

Theory of Cold Atoms: Basics of Quantum Statistics

V.I. Yukalov

*Bogolubov Laboratory of Theoretical Physics,
Joint Institute for Nuclear Research, Dubna 141980, Russia*

Abstract

The aim of this Tutorial is to present the basic mathematical techniques required for an accurate description of cold trapped atoms, both Bose and Fermi. The term *cold* implies that considered temperatures are low, such that quantum theory is necessary, even if temperatures are finite. And the term *atoms* means that the considered particles are structureless, being defined by their masses and mutual interactions. Atoms are *trapped* in the sense that they form a finite quantum system, though their number can be very large allowing for the use of the methods of statistical mechanics. This Tutorial is the first part of several tutorials, giving general mathematical techniques for both types of particle statistics. The following tutorials will be devoted separately to Bose atoms and Fermi atoms. The necessity of carefully explaining basic techniques is important for avoiding numerous misconceptions often propagating in literature.

PACS numbers: 03.70.+k; 03.75.-b; 05.30.Ch; 05.70.Ce; 05.70.Fh; 64.70.qd; 67.10.-j; 67.10.Fj

Keywords: Quantized fields; Cold atoms; Quantum ensembles; Thermodynamic functions; Phase transitions; Quantum matter; Quantum statistics

1 Introduction

Physics of cold atoms has recently been attracting high attention because of the breakthrough in experimental techniques of cooling and trapping atoms, resulting in a series of experiments with cold trapped atoms, when quantum effects become principal. In these experiments, different atoms can be prepared and studied in a number of states, with varying density, temperature, atomic interactions, and external fields. This allows for a detailed comparison of theory and experiment. However, in order that such a comparison would be meaningful, the used theory must be self-consistent and mathematically correct. In the other case, various misconceptions and paradoxes can arise. In theoretical literature, one can meet numerous incorrect statements and unjustified approximations, which especially concerns systems with Bose-Einstein condensate. It looks, therefore, that a Tutorial, explaining in detail the mathematical techniques that are necessary for a correct description of cold atoms is timely and important.

The present Tutorial is devoted to the general techniques of characterizing atomic systems for which quantum theory is necessary. Applications to Bose and Fermi systems are planned to be treated separately in the following tutorials. To make the presentation self-consistent, it is written with sufficient details, in order that the reader would not be forced to jump to other numerous publications. The related literature is cited. But, since the Tutorial is not a review, the citations are mainly limited to books, review articles, and most important original papers. Some topics, relevant to the physics of cold atoms have been described in the books [1–4] and review articles [5–22]. Several results, presented in this Tutorial are new.

Throughout the text, the system of units is employed, where the Planck constant $\hbar \rightarrow 1$ and the Boltzmann constant $k_B \rightarrow 1$.

2 Phenomenological relations

Before starting the exposition of a rigorous microscopic approach, it is useful to recall in brief the main phenomenological relations that often are common for quantum as well as classical systems and, when applied to quantum systems, describe the relations between their average characteristics.

2.1 Thermodynamic quantities

Thermodynamic relations are the same for either quantum or classical systems, provided that the considered system is in equilibrium and is sufficiently large, with its number of atoms N and volume V such that $N \gg 1$ and the volume being proportional to N .

The transition to bulk matter corresponds to the *thermodynamic limit*

$$N \rightarrow \infty, \quad V \rightarrow \infty, \quad \frac{N}{V} \rightarrow \text{const}, \quad (1)$$

so that the average atomic density $\rho \equiv N/V$ is finite, $0 < \rho < \infty$.

In the case of finite systems, it is not necessary to invoke the thermodynamic limit, though, for the validity of thermodynamic relations, the system is to be sufficiently large, as is mentioned above. When atoms are trapped in an external confining potential, instead of

limit (1), one defines [23] the *effective thermodynamic limit*

$$N \rightarrow \infty, \quad A_N \rightarrow \infty, \quad \frac{A_N}{N} \rightarrow \text{const} , \quad (2)$$

where A_N is an extensive observable quantity. If A_N is proportional to V , then the effective limit (2) reduces to the standard limit (1).

The notion of thermodynamic limit is important not only for bulk matter, but for finite systems as well, being closely related to system stability. Even dealing with finite systems, it is instructive to check the validity of thermodynamic relations under the effective thermodynamic limit, which makes it possible to understand whether the finite system is stable.

Thermodynamic quantities are distinguished into intensive and extensive. *Intensive* quantities, being defined for a large system, do not depend on asymptotically large N , remaining finite for $N \rightarrow \infty$. Examples of such intensive quantities are temperature T , pressure P , chemical potential μ , magnetic field \mathbf{B} , electric field \mathbf{E} , specific heat C_V , isothermal compressibility κ_T , and other susceptibilities.

Contrary to this, *extensive* quantities are proportional to N in some positive power, so that they tend to infinity together with N . For systems confined in a volume V , extensive quantities are proportional to V as well as to N . When the atomic system is trapped in a confining potential, and the volume is not defined, then extensive quantities can be proportional to N^α , with $\alpha \leq 1$. Examples of extensive quantities are energy E , entropy S , magnetic moment (magnetization) \mathbf{M} , electric moment (polarization) \mathbf{P} , and thermodynamic potentials.

Thermodynamic potentials define the thermodynamic state of the system. Each thermodynamic potential is a function of those thermodynamic variables that are necessary and sufficient for a unique description of thermodynamics of an *equilibrium system*. The most often used thermodynamic potentials are as follows.

Internal energy $E = E(S, V, N, \mathbf{M})$. In the presence of an electric field \mathbf{E} , the latter enters all equations similarly to the magnetic field \mathbf{B} , because of which we show below only the dependence on \mathbf{B} . In the same way, electric polarization \mathbf{P} enters similarly to magnetization \mathbf{M} . The internal energy differential is

$$dE = TdS - PdV + \mu dN + \mathbf{B} \cdot d\mathbf{M} . \quad (3)$$

Hence, one has

$$T = \frac{\partial E}{\partial S} , \quad P = - \frac{\partial E}{\partial V} , \quad \mu = \frac{\partial E}{\partial N} , \quad B^\alpha = \frac{\partial E}{\partial M^\alpha} , \quad (4)$$

where $\alpha = x, y, z$.

Free energy $F = F(T, V, N, \mathbf{M})$. It is related to the internal energy as

$$F = E - TS , \quad (5)$$

and its differential is

$$dF = -SdT - PdV + \mu dN + \mathbf{B} \cdot d\mathbf{M} . \quad (6)$$

Therefore, one has

$$S = - \frac{\partial F}{\partial T} , \quad P = - \frac{\partial F}{\partial V} , \quad \mu = \frac{\partial F}{\partial N} , \quad B^\alpha = \frac{\partial F}{\partial M^\alpha} .$$

Gibbs potential $G = G(T, P, N, \mathbf{B})$. It can be represented as

$$G = F + PV - \mathbf{B} \cdot \mathbf{M} = \mu N . \quad (7)$$

And its differential is

$$dG = -SdT + VdP + \mu dN - \mathbf{M} \cdot d\mathbf{B} \quad (8)$$

From here it follows:

$$S = -\frac{\partial G}{\partial T}, \quad V = \frac{\partial G}{\partial P}, \quad \mu = \frac{\partial G}{\partial N}, \quad M^\alpha = -\frac{\partial G}{\partial B^\alpha} .$$

Grand potential $\Omega = \Omega(T, V, \mu, \mathbf{M})$. It can be written as

$$\Omega = F - \mu N = -PV + \mathbf{B} \cdot \mathbf{M}, \quad (9)$$

with the differential

$$d\Omega = -SdT - PdV - Nd\mu + \mathbf{B} \cdot d\mathbf{M}. \quad (10)$$

Respectively,

$$S = -\frac{\partial \Omega}{\partial T}, \quad P = -\frac{\partial \Omega}{\partial V}, \quad N = -\frac{\partial \Omega}{\partial \mu}, \quad B^\alpha = \frac{\partial \Omega}{\partial M^\alpha} .$$

Some other thermodynamic potentials can be found in Kubo [24].

2.2 Equilibrium and stability

An equilibrium state of a system corresponds to an extremum of a thermodynamic potential, when

$$\delta E = \delta F = \delta G = \delta \Omega = 0, \quad (11)$$

with respect to the variation of parameters.

An equilibrium can be stable, metastable, or unstable. The absolute minimum of a thermodynamic potential corresponds to a thermodynamically absolutely stable system. A relative minimum represents a metastable system. And a maximum characterizes an unstable system. The sufficient conditions for the thermodynamic stability are

$$\delta^2 E > 0, \quad \delta^2 F > 0, \quad \delta^2 G > 0, \quad \delta^2 \Omega > 0. \quad (12)$$

These stability conditions can be expressed through second derivatives of thermodynamic potentials defining the related susceptibilities.

The system response to temperature variation is described by *specific heat* (heat capacity)

$$C_V \equiv \frac{T}{N} \left(\frac{\partial S}{\partial T} \right)_V. \quad (13)$$

Since $S = -(\partial F / \partial T)_V$, we have

$$C_V \equiv -\frac{T}{N} \left(\frac{\partial^2 F}{\partial T^2} \right)_V. \quad (14)$$

Using the relations $F = E - TS$ and

$$T \left(\frac{\partial S}{\partial T} \right)_V = \left(\frac{\partial E}{\partial T} \right)_V ,$$

we get

$$C_V = \frac{1}{N} \left(\frac{\partial E}{\partial T} \right)_V . \quad (15)$$

The *isothermal compressibility* describes the system response to pressure variation,

$$\kappa_T \equiv - \frac{1}{V} \left(\frac{\partial V}{\partial P} \right)_{TN} . \quad (16)$$

Substituting $V = (\partial G / \partial P)_{TN}$ gives

$$\kappa_T = - \frac{1}{V} \left(\frac{\partial^2 G}{\partial P^2} \right)_{TN} . \quad (17)$$

The isothermal compressibility can also be written as

$$\kappa_T = - \frac{1}{V} \left(\frac{\partial P}{\partial V} \right)_{TN}^{-1} = \frac{1}{V} \left(\frac{\partial^2 F}{\partial V^2} \right)_{TN}^{-1} . \quad (18)$$

And, with $\rho \equiv N/V$, one has

$$\kappa_T = \frac{1}{\rho} \left(\frac{\partial \rho}{\partial P} \right)_{TN} = \frac{1}{\rho} \left(\frac{\partial P}{\partial \rho} \right)_{TN}^{-1} . \quad (19)$$

The isothermal compressibility is a very important quantity characterizing the system thermodynamic stability, because of which it is useful to have its other representations. Let there be no magnetic and electric fields. Then the Gibbs potential differential is $dG = -SdT + VdP + \mu dN$. Since $G = \mu N$, it follows $Nd\mu = -SdT + VdP$, from where

$$\left(\frac{\partial \mu}{\partial N} \right)_{TV} = \frac{1}{\rho} \left(\frac{\partial P}{\partial N} \right)_{TV} .$$

Inverting this yields

$$\left(\frac{\partial N}{\partial \mu} \right)_{TV} = \rho \left(\frac{\partial N}{\partial P} \right)_{TV} .$$

For the related variables N, P, V , one can write

$$dN = \left(\frac{\partial N}{\partial P} \right)_V dP + \left(\frac{\partial N}{\partial V} \right)_P dV .$$

Under the fixed N , implying the zero total variation $dN = 0$, we have

$$\left(\frac{\partial N}{\partial P} \right)_V \left(\frac{\partial P}{\partial V} \right)_N + \left(\frac{\partial N}{\partial V} \right)_P = 0 ,$$

from where

$$\left(\frac{\partial N}{\partial P}\right)_V \left(\frac{\partial P}{\partial V}\right)_N \left(\frac{\partial V}{\partial N}\right)_P = -1.$$

Using the Maxwell relations

$$\left(\frac{\partial V}{\partial N}\right)_{TP} = \left(\frac{\partial \mu}{\partial P}\right)_{TN}$$

and the equalities

$$\left(\frac{\partial \mu}{\partial P}\right)_{TN} = \frac{1}{N} \left(\frac{\partial G}{\partial P}\right)_{TN} = \frac{1}{\rho},$$

we get

$$\left(\frac{\partial N}{\partial P}\right)_{TV} = -\rho \left(\frac{\partial V}{\partial P}\right)_{TN}.$$

Hence

$$\left(\frac{\partial V}{\partial P}\right)_{TN} = -\frac{1}{\rho} \left(\frac{\partial N}{\partial P}\right)_{TV} = -\frac{1}{\rho^2} \left(\frac{\partial N}{\partial \mu}\right)_{TV}.$$

This gives

$$\kappa_T \equiv \frac{1}{N} \left(\frac{\partial N}{\partial P}\right)_{TV} = \frac{1}{N\rho} \left(\frac{\partial N}{\partial \mu}\right)_{TV} = \frac{V}{\rho^2} \left(\frac{\partial^2 F}{\partial \rho^2}\right)_{TV}^{-1}. \quad (20)$$

The isothermal compressibility is closely connected with the sound velocity s , defined by the equation

$$s^2 \equiv \frac{1}{m} \left(\frac{\partial P}{\partial \rho}\right)_{TN}, \quad (21)$$

where m is atomic mass. As is clear,

$$\kappa_T = \frac{1}{m\rho s^2}. \quad (22)$$

Magnetic susceptibility is defined through the susceptibility tensor

$$\chi_{\alpha\beta} \equiv \frac{1}{N} \left(\frac{\partial M^\alpha}{\partial B^\beta}\right). \quad (23)$$

Using here $M^\alpha = -\partial G/\partial B^\alpha$ gives

$$\chi_{\alpha\beta} = -\frac{1}{N} \left(\frac{\partial^2 G}{\partial B^\alpha \partial B^\beta}\right). \quad (24)$$

Diagonal parts $\chi_{\alpha\alpha}$ of tensor (23) are called magnetic susceptibilities.

The thermodynamic characteristics, describing the system response to the variation of thermodynamic variables and which are expressed through the second derivatives of thermodynamic potentials, all can be termed susceptibilities. Thus, the specific heat C_V is *thermal susceptibility* (susceptibility with respect to temperature variation), the isothermal compressibility κ_T is *mechanical susceptibility* (susceptibility with respect to pressure variation), and the magnetic susceptibility $\chi_{\alpha\alpha}$ defines the system response to magnetic field variation. According to the stability conditions (12), all these susceptibilities are to be non-negative. Being intensive quantities, the susceptibilities of stable equilibrium systems

have to be finite, even in thermodynamic limit. In this way, the necessary conditions for *thermodynamic stability* of an equilibrium system are:

$$0 \leq C_V < \infty, \quad 0 \leq \kappa_T < \infty, \quad 0 \leq \chi_{\alpha\alpha} < \infty. \quad (25)$$

These stability conditions are the direct consequence of the convexity properties (12) of thermodynamic potentials and of the fact that susceptibilities are intensive quantities. The compulsory finiteness of the susceptibilities is easily understood. Take, for instance, specific heat C_V . If it would be infinite, then, according to Eq. (15), an infinitesimally weak temperature increase would yield infinite system energy, which is meaningless. If compressibility κ_T would be infinite, this, by definition (16), would imply that an infinitesimally small fluctuation of pressure would make the system volume either zero or infinity. Similar arguments are applicable to magnetic susceptibility.

2.3 Phases and symmetries

Thermodynamic phase, defined by a set of thermodynamic variables, is a system state qualitatively different from other states of the same system [24]. To make this vague definition more precise, one introduces *order parameters* such that to distinct phases there would correspond different orders. An order parameter is a quantity specifying a given thermodynamic phase and qualitatively distinguishing it from other phases.

For example, the distinction between ferromagnetic and paramagnetic phases is done by means of magnetization \mathbf{M} , so that $M \equiv |\mathbf{M}|$ is nonzero for ferromagnetic phase ($M > 0$, ferromagnet) and is zero for paramagnetic phase ($M = 0$, paramagnet). Similarly, the nonzero or zero polarization \mathbf{P} distinguishes ferroelectric from paraelectric phases.

Crystalline phase is characterized by a set of spatial vectors $\{\mathbf{a}\}$, called lattice vectors, with respect to which the particle density is periodic: $\rho(\mathbf{r} + \mathbf{a}) = \rho(\mathbf{r})$. While liquid-phase density is uniform: $\rho(\mathbf{r}) = \rho$. The role of an order parameter, distinguishing between the crystalline and liquid phases can be played by the difference

$$\Delta\rho \equiv \max_{\mathbf{r}} \rho(\mathbf{r}) - \min_{\mathbf{r}} \rho(\mathbf{r}).$$

For crystalline phase, $\Delta\rho > 0$, while for liquid phase, $\Delta\rho = 0$.

Different types of crystalline lattices are described by different sets of lattice vectors. For instance, one lattice, with the set $\{\mathbf{a}\}$ can be simple cubic, with the particle density $\rho_1(\mathbf{r} + \mathbf{a}) = \rho_1(\mathbf{r})$, while another lattice, with the set $\{\mathbf{b}\}$, can be face-centered cubic, having the density periodicity $\rho_2(\mathbf{r} + \mathbf{b}) = \rho_2(\mathbf{r})$.

Gaseous and liquid phases both are characterized by uniform density. Hence the latter cannot be an order parameter for distinguishing between these phases. The difference between liquid and gas can be noticed on the microscopic level by studying the pair correlation function displaying a kind of short-range order in the case of liquid. This type of short-range order can be captured by means of the order indices [25, 26].

Systems with Bose-Einstein condensate and without it can be distinguished by the related condensate density ρ_0 . When $\rho_0 > 0$, the system is Bose-condensed, and when $\rho_0 = 0$, it is not Bose-condensed.

Superfluid and normal fluid are distinguished by superfluid density ρ_s . The system is superfluid if $\rho_s > 0$, while it is normal if $\rho_s = 0$.

Dielectric and metallic phases can be distinguished by the single-electron energy gap ΔE , with $\Delta E > 0$ representing dielectric phase and $\Delta E = 0$ corresponding to metallic phase.

Superconductor and normal conductor are characterized by the presence or absence of a gap Δ in the spectrum of collective excitations. For a superconductor, $\Delta > 0$, while for a conductor, $\Delta = 0$.

A system can be described by several order parameters corresponding to different properties. For example, it can be ferromagnetic and dielectric or ferromagnetic and metallic.

Thermodynamic phases, having different order parameters (or order indices), are usually related to different types of symmetry. Under symmetry one understands invariance under some transformations forming a group. A symmetry group is defined as follows.

Let \hat{g}_i be a transformation labelled by an index i . And let the family of such transformations satisfy the additivity property $\hat{g}_i \cdot \hat{g}_j = \hat{g}_{i+j}$. If the family includes a unit element $\hat{g}_0 = \hat{1}$ and, for each \hat{g}_i there exists an inverse element, such that $\hat{g}_i^{-1} \hat{g}_i = \hat{1}$, then the family $\{\hat{g}_i\}$ forms a group.

As an example, consider the *inversion symmetry* that can be illustrated by inverting the magnetization vector \mathbf{M} of a ferromagnet. Denoting the inversion transformation by \hat{I} , we have $\hat{I}\mathbf{M} = -\mathbf{M}$. If the order parameter is given by $|\mathbf{M}|$, it is symmetric with respect to inversion: $\hat{I}M = |-\mathbf{M}| = |\mathbf{M}| = M$. Two elements, the inversion transformation and unity, form the inversion group $\{\hat{I}, \hat{1}\}$.

Another example is the *rotation symmetry*. Denote the spatial rotation transformation by \hat{R} . The order parameter M is invariant under the rotation group $\{\hat{R}\}$, since $\hat{R}M = |\hat{R}\mathbf{M}| = |\mathbf{M}| = M$.

One more example is the *translation symmetry*. Let $\{\hat{T}(\mathbf{r})\}$ be a translation transformation by a vector \mathbf{r} . If \mathbf{a} is a vector of a crystalline lattice, then the crystal density is invariant under the transformation

$$\hat{T}(\mathbf{a})\rho(\mathbf{r}) = \rho(\mathbf{r} + \mathbf{a}) = \rho(\mathbf{r}) .$$

The family $\{\hat{T}(\mathbf{a})\}$ composes the lattice translation group.

In the case of a liquid, one has $\hat{T}(\mathbf{r})\rho = \rho$ for any vector \mathbf{r} . Then the family $\{\hat{T}(\mathbf{r})\}$ forms the total translation group.

2.4 Phase transitions

A phase transition is a transformation of one thermodynamic phase into another. The change happens under a varying thermodynamic parameter, say, temperature T . When T reaches the stability boundary of one phase, the system becomes unstable and, under the following variation of T , transforms into a stable phase. The point of a phase transition is the point of instability. Therefore, at this point, some susceptibilities may diverge. Since thermodynamic phases can usually be characterized by order parameters, at the phase transition point, the order parameter qualitatively changes. Depending on the kind of this change, one distinguishes three types of phase transitions.

Let an order parameter be denoted by η , being in particular cases either magnetization M , or particle density ρ , or condensate density ρ_0 , or superfluid density ρ_s , or single-particle gap ΔE , or the gap in the spectrum of collective excitations Δ . And let temperature T be varying, with T_0 being the point of a phase transition. The latter are classified into three types.

(i) *First-order phase transition.* This is a discontinuous transition, with a finite jump of an order parameter η at the transition point:

$$|\eta(T_0 + 0) - \eta(T_0 - 0)| > 0 .$$

The values of $\eta(T_0 \pm 0)$ are not necessarily zero. An example of a first-order phase transition is the melting-crystallization transition between a crystal and a liquid, where there occurs a discontinuity of the density.

(ii) *Second-order phase transition.* This is a continuous transition, with a continuous variation of an order parameter between its non-zero and zero values:

$$|\eta(T_c + 0) - \eta(T_c - 0)| = 0 ,$$

where $\eta(T)$ becomes zero for one of the phases, say for $T \geq T_c$, being non-zero for the other phase. The point of a second-order phase transition is also called the *critical point*, since some of the susceptibilities diverge there. Respectively, the transition temperature is denoted as T_c , to distinguish it from the first-order transition point T_0 . Examples of second-order transitions are different magnetic transitions or Bose-Einstein condensation.

(iii) *Crossover transition.* This is a fast but gradual variation of an order parameter that is non-zero from both sides of the crossover point T_c . The crossover point can be defined as the *inflection point* of the order parameter:

$$\frac{\partial^2 \eta}{\partial T^2} = 0 \quad (T = T_c) .$$

That is, the first derivative of the order parameter displays an extremum at the inflection point.

In the vicinity of a phase-transition point, the system properties may essentially change, because of which a region around a transition point, where such changes are strong, is termed the *critical region*.

If thermodynamic phases are described by different symmetries, the latter also change at the phase transition point. When the symmetry of one of the phases is lower than that of the other phase, the transition to the lower-symmetry phase is called *spontaneous symmetry breaking*. It is termed *spontaneous*, since the symmetry breaking occurs in a self-organized way, without action of external fields. The cause for the spontaneous symmetry breaking is that the phase with a lower symmetry happens to be stable, when the symmetric phase becomes unstable.

More information on phase transitions can be found, e.g., in the books [27–32].

2.5 Heterophase mixtures

In addition to pure thermodynamic phases, there can arise mixtures of different phases, where the latter are not spatially completely separated, but are intermixed in such a way that each of the phases occupies randomly distributed in space regions of *mesoscopic* size. Such mesoscopic mixtures are also called *heterophase mixtures* [33]. For example, we can imagine paramagnetic regions in a ferromagnet, ferromagnetic clusters in a paramagnet, disordered regions in a crystal, crystalline clusters in a liquid, gas bubbles in a liquid, and liquid droplets in a gas.

For illustration, let us keep in mind the case of a heterophase mixture of two phases, characterized by different order parameters, η_1 and η_2 . When the system is composed of two macroscopic thermodynamic phases, they are said to be in thermal equilibrium with each other if their temperatures coincide, $T_1 = T_2$. The phases are in mechanical equilibrium, if their pressures are the same, $P_1 = P_2$. And they are in chemical equilibrium, if their chemical potentials are equal, $\mu_1 = \mu_2$. A mesoscopic heterophase mixture, strictly speaking, is quasi-equilibrium, because of which the above equilibrium conditions are valid only on average [33].

Let V_ν be the volume occupied by a ν -phase and N_ν be the number of particles in the ν -phase. Mesoscopic phases can be separated by means of an *equimolecular separating surface* [34, 35]. In that case, all extensive quantities are additive, e.g.,

$$V_1 + V_2 = V, \quad N_1 + N_2 = N.$$

The geometric weights of each phase are quantified by

$$w_\nu \equiv \frac{V_\nu}{V} \quad (\nu = 1, 2).$$

This quantity has the meaning of geometric phase probability and satisfies the conditions

$$0 \leq w_\nu \leq 1, \quad w_1 + w_2 = 1.$$

Equations for the phase probabilities are obtained from the minimization of a thermodynamic potential under the above normalization condition. Defining, for convenience,

$$w_1 \equiv w, \quad w_2 \equiv 1 - w,$$

one has to find the absolute minimum of a thermodynamic potential, say, of the free energy $F = F(w)$:

$$F = \text{abs} \min_w F(w).$$

The minimization has to include the consideration of w at the boundary, where either $w = 0$ or $w = 1$. That is, it is necessary to compare three quantities: $F(w)$, with $0 < w < 1$, and the boundary values $F(0)$, and $F(1)$.

The condition of heterophase equilibrium reads as

$$\frac{\partial F}{\partial w} = 0.$$

And the condition of heterophase stability is

$$\frac{\partial^2 F}{\partial w^2} > 0.$$

These equations define the geometric phase probabilities w_ν playing the role of additional order parameters. The second derivative $\partial^2 F / \partial w^2$ has the meaning of the *heterophase susceptibility* that is required to be non-negative and finite for guaranteeing the system stability, similarly to stability conditions (25).

2.6 Landau potential

If the considered phase transition is of second order or of weak first order, when the order parameter in the vicinity of the transition point is small, it is possible, assuming that the free energy is an analytical function of the order parameter, to expand the free energy in powers of the latter, as suggested by Landau [36].

Let us consider the critical region of temperatures $T \approx T_c$, where the order parameter $\eta = \eta(T)$ is small. The order parameter η is assumed to be real-valued and such that $\eta(T) \neq 0$ for $T < T_c$, but $\eta(T) = 0$ for $T > T_c$. The free energy $F = F(\eta)$ is expanded in powers of the order parameter as

$$F = F(0) + a_1\eta + a_2\eta^2 + a_3\eta^3 + a_4\eta^4 .$$

To have a minimum at finite η , it is necessary that $a_4 > 0$.

In equilibrium, one has

$$\frac{\partial F}{\partial \eta} = a_1 + 2a_2\eta + 3a_3\eta^2 + 4a_4\eta^3 = 0 ,$$

which requires that $a_1 = 0$, when $\eta = 0$. The condition of stability is

$$\frac{\partial^2 F}{\partial \eta^2} = 2(a_2 + 3a_3\eta + 6a_4\eta^2) \geq 0 .$$

There can happen two cases, with and without inversion symmetry.

In the case of *inversion symmetry*, the free energy is such that $F(-\eta) = F(\eta)$. Then $a_3 = 0$ and the condition of equilibrium reads as

$$\eta(a_2 + 2a_4\eta^2) = 0 .$$

From here, we see that the order parameter at low temperatures can take one of the two values

$$\eta_{\pm} = \pm \sqrt{-\frac{a_2}{2a_4}} \quad (T < T_c) .$$

And above T_c , one has

$$\eta = 0 \quad (T > T_c) .$$

The stability condition yields

$$a_2 + 6a_4\eta^2 \geq 0 ,$$

which implies that

$$\begin{aligned} a_2(T) &< 0 & (T < T_c) , \\ a_2(T) &\geq 0 & (T > T_c) . \end{aligned}$$

This suggests that one can accept the form

$$a_2(T) = \alpha(T - T_c) .$$

The free energy is symmetric with respect to the inversion of η to $-\eta$. The two order parameters have the same magnitude, $|\eta_+| = |\eta_-|$. This situation corresponds to *macroscopic*

degeneracy. When temperature is lowered, the system passes from the symmetric phase, with $\eta = 0$, to the phase with a non-zero value of the order parameter, taking spontaneously either η_+ or η_- . This is an example of spontaneous symmetry breaking.

When there is *no inversion symmetry*, then $a_3 \neq 0$, and the condition of equilibrium defines the order parameter as

$$\eta_{\pm} = \frac{-3a_3 \pm \sqrt{9a_3^2 - 32a_2a_4}}{8a_4} .$$

To have a real-valued order parameter, it is necessary that

$$9a_3^2 - 32a_2a_4 \geq 0 .$$

This situation corresponds to a first-order phase transition.

2.7 Ginzburg-Landau functional

In the Landau expansion, the order parameter is treated as a scalar. In general, it can be an order function $\eta = \eta(\mathbf{r}, t)$ of the real-space vector \mathbf{r} and time t . The generalization to this case has been proposed by Ginzburg and Landau [37].

One defines an effective Hamiltonian density $H(\mathbf{r}) = H(\eta(\mathbf{r}, t), T, \dots)$ depending on the order function and thermodynamic variables, such as temperature T . The effective Hamiltonian

$$H = \int H(\mathbf{r}) d\mathbf{r}$$

depends on T , reminding rather free energy. This effective Hamiltonian is the *Ginzburg-Landau functional* of the order function.

The dependence of the order function on the spatial and temporal variables implies that the system is not uniform and, in general, not equilibrium. The evolution equation for the order function is

$$i \frac{\partial \eta}{\partial t} = \frac{\delta H}{\delta \eta^*} . \quad (26)$$

Stationary states can be represented as

$$\eta(\mathbf{r}, t) = \psi(\mathbf{r}) e^{-i\varepsilon t} ,$$

where ε is a real-valued quantity. Then the evolution equation reduces to the stationary form

$$\varepsilon \psi = \frac{\delta H}{\delta \psi^*} .$$

Introducing the grand effective Hamiltonian, or grand potential

$$\Omega \equiv H - \varepsilon \int |\psi|^2 d\mathbf{r} ,$$

the stationary equation can be rewritten as

$$\frac{\delta \Omega}{\delta \psi^*} = 0 ,$$

which is a condition of equilibrium.

The function ψ can be normalized to the number of particles N by the normalization condition

$$\int |\psi|^2 d\mathbf{r} = N, \quad N = -\frac{\partial \Omega}{\partial \varepsilon}.$$

This shows that ε plays the role of a chemical potential.

The dynamic stability of the stationary solution is checked by studying small deviations from this solution. For this purpose, the order function is written as

$$\eta(\mathbf{r}, t) = [\psi(\mathbf{r}) + \delta\psi(\mathbf{r}, t)]e^{-i\varepsilon t},$$

with the deviation

$$\delta\psi(\mathbf{r}, t) = u(\mathbf{r})e^{-i\omega t} + v^*(\mathbf{r})e^{i\omega t}$$

representing the wave function of collective excitations. Substituting the above representation into the evolution equation yields an equation for ω . The condition of *dynamic stability* requires that the admissible ω be positive (non-negative). This is because if $\omega = \omega_1 + i\omega_2$ is complex-valued, then

$$e^{-i\omega t} = e^{-i\omega_1 t}e^{\omega_2 t}, \quad e^{i\omega t} = e^{i\omega_1 t}e^{-\omega_2 t},$$

and there appears a divergence for any $\omega_2 \neq 0$, as time increases, which means that the considered solution is not dynamically stable.

2.8 Superconductor model

Ginzburg and Landau [37] illustrated their approach by a model of superconductor. In the effective Hamiltonian $H = \int H(\mathbf{r}) d\mathbf{r}$, the Hamiltonian density is written as

$$H(\mathbf{r}) = \frac{1}{2m} \left| \left(\hat{\mathbf{p}} - \frac{e}{c} \mathbf{A} \right) \eta \right|^2 + U|\eta|^2 + \frac{1}{2} \Phi_0 |\eta|^4 + \frac{\mathbf{B}^2}{8\pi},$$

where m is the effective mass equal to two electron masses, $m = 2m_0$, and the effective charge equals two electron charges, $e = -2e_0$; momentum operator $\hat{\mathbf{p}} \equiv -i\nabla$, and \mathbf{A} is the vector potential. The quantities U and Φ_0 can be functions of thermodynamic variables T, P, \dots . The order parameter η is such that $|\eta|^2$ gives the superfluid electron density.

In the presence of an external magnetic field \mathbf{B}_0 , the effective Hamiltonian becomes

$$H = \int \left[H(\mathbf{r}) - \frac{\mathbf{B} \cdot \mathbf{B}_0}{4\pi} \right] d\mathbf{r}.$$

The vector potential defines the magnetic field $\mathbf{B} = \nabla \times \mathbf{A}$ and satisfies the Coulomb calibration $\nabla \cdot \mathbf{A} = 0$. In equilibrium, one has

$$\frac{\delta H}{\delta \eta^*(\mathbf{r})} = 0.$$

This leads to the *equation for the order function*

$$\frac{1}{2m} \left(\hat{\mathbf{p}} - \frac{e}{c} \mathbf{A} \right)^2 \eta + U\eta + \Phi_0 |\eta|^2 \eta = 0. \quad (27)$$

The extremization with respect to the vector potential,

$$\frac{\delta H}{\delta A^\alpha} = 0 ,$$

can be rewritten as the Euler-Lagrange equations

$$\left(\frac{\delta H}{\delta A^\alpha} \right)_{A'} - \sum_{\beta} \frac{\partial}{\partial r^\beta} \left(\frac{\delta H}{\delta A'_{\alpha\beta}} \right) = 0 ,$$

in which $A'_{\alpha\beta} \equiv \partial A^\alpha / \partial r^\beta$. From here, one finds the equation

$$\nabla^2 \mathbf{A} = -\frac{4\pi}{c} \mathbf{J} ,$$

with the supercurrent

$$\mathbf{J} = -\frac{ie}{2m} (\eta^* \nabla \eta - \eta \nabla \eta^*) - \frac{e^2}{mc} |\eta|^2 \mathbf{A} .$$

The equation for the order function η is a kind of the nonlinear Schrödinger equation. It is worth emphasizing that the order-function equation, being a kind of the nonlinear Schrödinger equation, has appeared, first, in the paper by Ginzburg and Landau [37]. Therefore such equations can be called the *Ginzburg-Landau order-function equations*.

2.9 Superfluid vortex

For systems of neutral Bose atoms, such as superfluid helium, the effective Hamiltonian density takes the form

$$H(\mathbf{r}) = \frac{1}{2m} |\nabla \eta|^2 + U |\eta|^2 + \frac{1}{2} \Phi_0 |\eta|^4 ,$$

where m is atomic mass and the quantities U and Φ_0 can be functions of thermodynamic variables T, \dots . The order parameter, or order function $\eta = \eta(\mathbf{r}, t)$ gives the density $|\eta|^2$ of atoms in the ordered state.

With the nonlinear Hamiltonian

$$\hat{H}[\eta] \equiv -\frac{\nabla^2}{2m} + U + \Phi_0 |\eta|^2 , \quad (28)$$

the evolution equation for the order function

$$i \frac{\partial \eta}{\partial t} = \hat{H}[\eta] \eta \quad (29)$$

is the nonlinear Schrödinger equation, being a particular case of the Ginzburg-Landau equation (27). For an equilibrium system, one gets the stationary equation

$$\hat{H}[\psi] \psi = \varepsilon \psi . \quad (30)$$

The latter equation possesses vortex solutions [38]. To show this, one passes to cylindrical coordinates $\mathbf{r} = \{r_\perp, \varphi, z\}$, where $r_\perp \equiv \sqrt{x^2 + y^2}$ is the transverse spatial variable. In this representation,

$$\nabla^2 = \frac{\partial^2}{\partial r_\perp^2} + \frac{1}{r_\perp} \frac{\partial}{\partial r_\perp} + \frac{1}{r_\perp^2} \frac{\partial^2}{\partial \varphi^2} + \frac{\partial^2}{\partial z^2} .$$

Looking for the solution in the form $\psi(\mathbf{r}) = \psi_\perp(r_\perp)e^{i\nu\varphi}$, with $\psi_\perp(r_\perp)$ real-valued, and imposing the periodicity constraint

$$\psi(r_\perp, \varphi + 2\pi, z) = \psi(r_\perp, \varphi, z) ,$$

one finds that the quantum number ν , called the *winding number*, has to be quantized, $\nu = 0, \pm 1, \pm 2, \dots$

The stationary equation (30) yields

$$\left[\frac{d^2}{dr_\perp^2} + \frac{1}{r_\perp} \frac{d}{dr_\perp} - \frac{\nu^2}{r_\perp^2} + 2m(\varepsilon - U) \right] \psi_\perp - 2m\Phi_0 \psi_\perp^3 = 0 . \quad (31)$$

Considering a cylindrical system of radius R , length L , and volume $V = \pi R^2 L$, the solution to this equation is normalized as

$$2\pi L \int_0^R |\psi_\perp(r_\perp)|^2 r_\perp dr_\perp = N . \quad (32)$$

One introduces the *healing length*

$$\xi \equiv \frac{1}{\sqrt{2m(\varepsilon - U)}} . \quad (33)$$

Using this scale, one defines the dimensionless radius and the dimensionless order function,

$$r \equiv \frac{r_\perp}{\xi} , \quad f(r) \equiv \xi \sqrt{2m\Phi_0} \psi_\perp(r_\perp) . \quad (34)$$

Then Eq. (31) reduces to

$$\left(\frac{d^2}{dr^2} + \frac{1}{r} \frac{d}{dr} + 1 - \frac{\nu^2}{r^2} \right) f - f^3 = 0 . \quad (35)$$

And the normalization condition (32) becomes

$$\frac{\pi L}{m\Phi_0} \int_0^{R/\xi} f^2(r) r dr = N . \quad (36)$$

Looking for a nontrivial solution $f \neq 0$, we see that it depends on the winding number ν .

If $\nu = 0$, there is no vortex, which corresponds to the *ground state*, when $f_0 = \text{const}$. From Eq. (35), we have $f_0^2 = 1$, hence $f_0 = \pm 1$. This degeneracy is not important, since the physical quantity $|f_0|^2$ does not depend on the sign of f_0 . Without the loss of generality, we may choose $f_0 = 1$. Normalization (36) gives

$$\xi = \frac{1}{\sqrt{2m\rho\Phi_0}} , \quad \varepsilon = U + \rho\Phi_0 \quad (\nu = 0) . \quad (37)$$

The first vortex solution appears for $\nu = 1$, describing the *basic vortex*. Then Eq. (35) can be solved numerically. Close to the vortex axis, where $r \ll 1$, that is $r_\perp \ll \xi$, the solution has the asymptotic expansion

$$f(r) \simeq \sqrt{C} r (1 + c_2 r^2 + c_4 r^4 + c_6 r^6) ,$$

with C being a constant and

$$c_2 = -\frac{1}{8} , \quad c_4 = \frac{1+8C}{192} , \quad c_6 = -\frac{1+80C}{9216} .$$

At large distance from the vortex axis, where $r \gg 1$, that is $r_\perp \gg \xi$, the solution has the form

$$f(r) \simeq 1 - \frac{1}{2} r^{-2} - \frac{9}{8} r^{-4} - \frac{161}{16} r^{-6} .$$

Solutions $f = f_\nu(r)$, representing the vortices with higher winding numbers, satisfy the equation

$$\left(\frac{d^2}{dr^2} + \frac{1}{r} \frac{d}{dr} + 1 - \frac{\nu^2}{r^2} \right) f_\nu - f_\nu^3 = 0 .$$

The normalization condition (36) defines the related healing length ξ_ν that is different from (37). Notation (32) gives the spectrum

$$\varepsilon_\nu = U + \frac{1}{2m\xi_\nu^2} , \tag{38}$$

labelled by the winding number $\nu = 1, 2, \dots$

If the system radius is very large, then the dependence of the healing length and, respectively, of the spectrum on the winding number is weak, since, in that case, the main contribution to the normalization integral (36) is due to the large region of $r \gg 1$, where $f_\nu(r) \rightarrow 1$. However for an elongated system, with a small radius, spectrum (38) can essentially depend on the winding number, if the vortex axis is along the cylinder axis.

The solution $\psi(\mathbf{r}) = \psi_\perp(r_\perp)e^{i\nu\varphi}$, for $\nu > 0$, represents a vortex, since it defines a nonzero tangential velocity

$$\mathbf{v} \equiv \frac{\mathbf{j}}{\rho} = \frac{\nu}{mr} \mathbf{e}_\varphi , \tag{39}$$

where

$$\rho \equiv |\psi|^2 , \quad \mathbf{j} = -\frac{i}{2m} (\psi^* \nabla \psi - \psi \nabla \psi^*) .$$

The vortex circulation corresponds to the integral

$$\oint \mathbf{v} \cdot d\mathbf{l} = 2\pi \frac{\nu}{m} . \tag{40}$$

It is worth emphasizing again that the evolution equations for order functions naturally appear in the Ginzburg-Landau approach in a similar way for both superconducting and superfluid systems [37–40].

2.10 Landau criterion

In order to find out whether the system can exhibit superfluid motion, one employs a simple criterion, suggested by Landau [36].

Suppose that a flow moves with superfluid velocity \mathbf{v} . In the moving frame, where the liquid is at rest, the energy per particle is

$$E = U + \frac{p^2}{2m} ,$$

with $p \equiv |\mathbf{p}|$. Its energy in the laboratory frame reads as

$$E_v = U + \frac{(\mathbf{p} + m\mathbf{v})^2}{2m} = U + \frac{p^2}{2m} + \mathbf{p} \cdot \mathbf{v} + \frac{mv^2}{2} ,$$

with $v \equiv |\mathbf{v}|$. If the energy of an elementary excitation in the moving frame is $\varepsilon(\mathbf{p})$, then its energy in the laboratory frame is $\varepsilon_v(\mathbf{p}) = \varepsilon(\mathbf{p}) + \mathbf{p} \cdot \mathbf{v}$.

The fluid motion is dynamically stable when the energy of the elementary excitations is positive,

$$\varepsilon_v(\mathbf{p}) > 0 , \quad \varepsilon(\mathbf{p}) + \mathbf{p} \cdot \mathbf{v} > 0 ,$$

hence, when $-\mathbf{p} \cdot \mathbf{v} < \varepsilon(\mathbf{p})$ for all \mathbf{p} . Maximizing the left-hand side of the latter inequality with respect to the direction of \mathbf{p} , one gets $pv < \varepsilon(\mathbf{p})$. Therefore, the superfluid velocity should be such that

$$v < \min_p \frac{\varepsilon(\mathbf{p})}{p} .$$

That is, it has to be lower than the critical velocity

$$v_c \equiv \min_p \frac{\varepsilon(\mathbf{p})}{p} .$$

For $v < v_c$, the superfluid motion is stable, while if v exceeds v_c , the superfluidity is destroyed.

Strictly speaking, superfluid flow is not absolutely stable, but metastable. The absolutely stable state corresponds to the system at rest.

As an example, when there is no superfluidity, it is possible to mention the ideal Bose gas with the spectrum $\varepsilon(\mathbf{p}) = p^2/2m$, for which $\min_p \varepsilon(\mathbf{p})/p = 0$, hence $v_c = 0$, and there can be no superfluid motion in the ideal gas, even as a metastable state.

3 Field representation

Now the basic mathematical notions will be given, providing the fundamentals for the microscopic theory of quantum systems. The first such a notion is the so-called second quantization. Below, we shall describe it by employing the field representation, following mainly Berezin [41] and Schweber [42].

3.1 Fock space

Let us enumerate the number of particles by $n = 0, 1, 2, \dots$. Each particle is characterized by a set x of variables from the variable space $\{x\}$. For the general mathematical exposition, the nature of the particle variables is not important. In particular, this can be the real-space vector \mathbf{r} , or momentum vector \mathbf{p} , or the combination of the real-space vector and a spin $\{\mathbf{r}, s\}$, etc.

The system microscopic states are the vectors of a Hilbert space. A *Hilbert space* \mathcal{H} is a complete vector space with a scalar product. In this space $\mathcal{H} = \{f\}$, for any two vectors $f, f' \in \mathcal{H}$ there exists a scalar product (f, f') . The latter generates the vector norm $\|f\| \equiv \sqrt{(f, f)}$. The space is complete, if for a sequence $\{f^{(k)} \in \mathcal{H} \mid k = 1, 2, \dots\}$, the limit $\lim_{k \rightarrow \infty} f^{(k)} = f^{(\infty)} \in \mathcal{H}$. The convergence is understood as the convergence by norm: $\lim_{k \rightarrow \infty} \|f^{(k)}\| = \|f^{(\infty)}\|$.

Let us consider the space $\mathcal{H}_n = \{f_n\}$, which is an n -particle *Hilbert space* composed of functions $f_n = f_n(x_1, x_2, \dots, x_n)$. Let \hat{P}_{ij} be a permutation operator interchanging the indices i and j , with $i \neq j$, so that its action is defined by the equality

$$\hat{P}_{ij} f_n(\dots x_i \dots x_j \dots) = f_n(\dots x_j \dots x_i \dots) .$$

It is assumed that all particles are classified into two types, bosons and fermions. The first type is described by symmetric states and the second, by antisymmetric. An n - particle function f_n is symmetrized provided that

$$\hat{P}_{ij} f_n = \pm f_n .$$

In what follows, the upper sign corresponds to bosons, while the lower sign, to fermions. The scalar product in \mathcal{H}_n is given by the integral

$$(f_n, f'_n) \equiv \int f_n^*(x_1, \dots, x_n) f'_n(x_1, \dots, x_n) dx_1 \dots dx_n .$$

The integration for each $x_i \in \{x_i\}$ is over all $\{x_i\}$, so that $\int dx \equiv \int_{\{x\}} dx$. Each space \mathcal{H}_n is complete by norm $\|f_n\| \equiv \sqrt{(f_n, f_n)}$. In the case of discrete variables, the integration is to be understood as summation. For instance, if $x \rightarrow \mathbf{r}, s$, where s is a spin index, then $\int dx \equiv \sum_s \int d\mathbf{r}$. In that case, the multi-dimensional delta function is $\delta(x - x') \equiv \delta(\mathbf{r} - \mathbf{r}') \delta_{ss'}$.

Consider a set of n -particle spaces $\mathcal{H}_0, \mathcal{H}_1, \dots$, where the first space \mathcal{H}_0 is composed of constants $f_0 = \text{const}$.

The *Fock space* is the direct sum

$$\mathcal{F} \equiv \bigoplus_{n=0}^{\infty} \mathcal{H}_n .$$

The vectors of the space $\mathcal{F} = \{\varphi\}$ have the form of columns

$$\varphi \equiv \begin{bmatrix} f_0 \\ f_1 \\ \cdot \\ \cdot \\ \cdot \\ f_n \\ \cdot \\ \cdot \\ \cdot \end{bmatrix}, \quad f_n = f_n(x_1, \dots, x_n).$$

The conjugate of φ is the row $\varphi^+ \equiv [f_0^* f_1^* \dots f_n^* \dots]$. For brevity, we shall denote $\varphi \equiv [f_n]$ and $\varphi^+ \equiv [f_n]^+$. The scalar product in the Fock space is

$$(\varphi, \varphi') \equiv \varphi^+ \varphi' \equiv \sum_{n=0}^{\infty} (f_n, f'_n)$$

for any $\varphi, \varphi' \in \mathcal{F}$. Hence, the Fock space is a Hilbert space. The norm is generated by the scalar product, $\|\varphi\| \equiv \sqrt{(\varphi, \varphi)}$.

Recall that functions f_n are assumed to be symmetrized, which can be done by means of the *symmetrization operator* \hat{S} that makes of functions either symmetric or antisymmetric combinations, depending on whether bosons or fermions are considered. Thus, a function of two variables is symmetrized as

$$\hat{S}h(x_i, x_j) \equiv \frac{1}{2}[h(x_i, x_j) \pm h(x_j, x_i)].$$

Similarly, functions of many variables are symmetrized.

Sometimes, one assumes that the symmetrization is done over the products of the type $f(x_1) \dots f(x_n)$. This, however, is not necessary, and the functions $f_n(x_1, \dots, x_n)$ can be of any type, provided they are symmetrized.

The Fock space \mathcal{F} , containing only symmetrized functions f_n is called the *ideal Fock space*.

The *vacuum* in the Fock space is the column

$$|0\rangle \equiv \begin{bmatrix} 1 \\ 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \end{bmatrix}.$$

The *restricted Fock space* \mathcal{F}_N for the fixed number of particles N consists of $\varphi_N \in \mathcal{F}_N$ such that $\varphi_N = [\delta_{nN} f_N]$. The restricted Fock space is the number-conserving space. This restriction makes operations in the Fock space more complicated, because of which one prefers to deal with the non-restricted Fock space. In what follows, we shall deal with the ideal Fock space that, for short, will be called just the Fock space.

3.2 Field operators

The field operators act on the Fock space. One terms $\psi^\dagger(x)$ the creation operator and $\psi(x)$, annihilation operator. The action of the creation operator is given by the formula

$$\psi^\dagger(x)\varphi \equiv \begin{bmatrix} 0 \\ f_0\delta(x_1 - x) \\ \sqrt{2} \hat{S}f(x_1)\delta(x_2 - x) \\ \cdot \\ \cdot \\ \cdot \\ \sqrt{n} \hat{S}f_{n-1}(x_1, \dots, x_{n-1})\delta(x_n - x) \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} .$$

In what follows, the latter form will be represented as

$$\psi^\dagger(x)\varphi = \left[\sqrt{n} \hat{S}f_{n-1}(x_1, \dots, x_{n-1})\delta(x_n - x) \right] .$$

The annihilation operator acts according to the rule

$$\psi(x)\varphi \equiv \begin{bmatrix} f_1(x) \\ \sqrt{2} f_2(x_1, x) \\ \sqrt{3} f_3(x_1, x_2, x) \\ \cdot \\ \cdot \\ \cdot \\ \sqrt{n+1} f_{n+1}(x_1, \dots, x_n, x) \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} .$$

This will be denoted as

$$\psi(x)\varphi = \left[\sqrt{n+1} f_{n+1}(x_1, \dots, x_n, x) \right] .$$

The field operators act on the vectors of the Fock space, but $\psi^\dagger(x)$ and $\psi(x)$ are not defined in the Fock space, since $\psi^\dagger(x)\varphi$ lies outside \mathcal{F} . In that sense, the field operators are the operator distributions.

In the Fock space, one can define the operators

$$\psi_f^\dagger \equiv \int f(x)\psi^\dagger(x) dx , \quad \psi_f \equiv \int f^*(x)\psi(x) dx ,$$

where $\|f\| < \infty$. Their action is given by the expressions

$$\psi_f^\dagger\varphi = \left[\sqrt{n} \hat{S}f_{n-1}(x_1, \dots, x_{n-1})f(x_n) \right] ,$$

$$\psi_f \varphi = \left[\sqrt{n+1} \int f_{n+1}(x_1, \dots, x_n, x) f^*(x) dx \right] ,$$

which are in \mathcal{F} .

The action of the annihilation operator on the vacuum is

$$\psi(x)|0\rangle = 0 .$$

And the creation operator acts on the vacuum as

$$\psi^\dagger(x)|0\rangle = \begin{bmatrix} 0 \\ \delta(x_1 - x) \\ 0 \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} .$$

A double action of the creation operators yields

$$\psi^\dagger(x'_1)\psi^\dagger(x'_2)|0\rangle = \begin{bmatrix} 0 \\ 0 \\ \sqrt{2} \hat{S} \delta(x_1 - x'_1) \delta(x_2 - x'_2) \\ 0 \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} .$$

Acting on the vacuum by n creation operators gives

$$\prod_{i=1}^n \psi^\dagger(x'_i)|0\rangle = \begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ \sqrt{n!} \hat{S} \prod_{i=1}^n \delta(x_i - x'_i) \\ 0 \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} .$$

Therefore, any $\varphi \in \mathcal{F}$ can be created by the creation operators by the formula

$$\varphi = \sum_{n=0}^{\infty} \frac{1}{\sqrt{n!}} \int f_n(x_1, \dots, x_n) \prod_{i=1}^n \psi^\dagger(x_i) dx_i |0\rangle .$$

The double action of the annihilation operators gives

$$\psi(x)\psi(x')\varphi = \left[\sqrt{(n+1)(n+2)} f_{n+2}(x_1, \dots, x_n, x, x') \right] .$$

This should be compared with the double action of the creation operators

$$\psi^\dagger(x)\psi^\dagger(x')\varphi = \left[\sqrt{n(n-1)} \hat{S} f_{n-2}(x_1, \dots, x_{n-2}) \delta(x_{n-1} - x) \delta(x_n - x') \right] .$$

Interchanging the variables x and x' , we come to the relations

$$\begin{aligned} \psi(x')\psi(x)\varphi &= \pm \psi(x)\psi(x')\varphi , \\ \psi^\dagger(x')\psi^\dagger(x)\varphi &= \pm \psi^\dagger(x)\psi^\dagger(x')\varphi . \end{aligned}$$

Since these equations are valid for any $\varphi \in \mathcal{F}$, we have

$$[\psi(x), \psi(x')]_{\mp} = [\psi^\dagger(x), \psi^\dagger(x')]_{\mp} = 0 .$$

Here the notation for the commutator and anticommutator of two operators is used, $[A, B]_{\mp} \equiv AB \mp BA$.

Similarly, we find

$$\begin{aligned} \psi(x)\psi^\dagger(x')\varphi &= \left[\pm(n+1) \hat{S} f_n(x_1, \dots, x_{n-1}, x) \delta(x_n - x') \right] , \\ \psi^\dagger(x')\psi(x)\varphi &= \left[n \hat{S} f_n(x_1, \dots, x_{n-1}, x) \delta(x_n - x') \right] . \end{aligned}$$

From here,

$$\begin{aligned} \psi_f \psi_g^\dagger \varphi &= [(n+1) f_n](f, g) , \\ \psi_g^\dagger \psi_f \varphi &= [\pm n f_n](f, g) . \end{aligned}$$

This can be rewritten as

$$[\psi_f, \psi_g^\dagger]_{\mp} \varphi = (f, g) \varphi , \quad [\psi_f, \psi_g^\dagger]_{\mp} = (f, g) .$$

Since these equations are valid for any $\varphi \in \mathcal{F}$ and any $f, g \in \mathcal{H}_1$, we come to the commutation (anticommutation) relations

$$[\psi(x), \psi^\dagger(x')]_{\mp} = \delta(x - x') .$$

A vector $\varphi \in \mathcal{F}$ is called a quantum state or microstate. The Fock space \mathcal{F} is a space of quantum states or a space of microstates. As is shown, the field operators $\psi(x)$ and $\psi^\dagger(x)$ act on the Fock space and can create any vector $[f_n(x_1, \dots, x_n)]$ in this space. Therefore, one tells that the Fock space is generated by ψ and denote this Fock space as $\mathcal{F}(\psi)$.

From the variables x , one can pass to any other required variables, for instance, changing the representation from x to k . For this purpose, one defines a complete orthonormal basis $\{\varphi_k(x)\}$ and expands the field operators over this basis,

$$\psi(x) = \sum_k a_k \varphi_k(x) , \quad \psi^\dagger(x) = \sum_k a_k^\dagger \varphi_k^*(x) .$$

The operators a_k and a_k^\dagger generate the Fock space $\mathcal{F}(a_k)$. The vectors of the latter space have the form $[h_n(k_1, \dots, k_n)] \in \mathcal{F}(a_k)$.

3.3 Operators of observables

In quantum mechanics, observable quantities are represented by operators of observables. An n -particle operator is written as

$$A_n = \sum_{i=1}^n A_1(x_i) + \frac{1}{2} \sum_{i \neq j}^n A_2(x_i, x_j) + \dots$$

It is a Hermitian or self-adjoint operator, such that $A_n^\dagger = A_n$. The kernels $A_n(x_1, \dots, x_n)$ are quantum-mechanical operators that are also self-adjoint. Examples of the quantum-mechanical operators are the operator of momentum

$$\mathbf{P}_n = \sum_{i=1}^n (-i \nabla_i),$$

where $\nabla_i \equiv \partial/\partial \mathbf{r}_i$, the kinetic-energy operator

$$K_n = \sum_{i=1}^n \left(-\frac{\nabla_i^2}{2m} \right),$$

and the potential-energy operator

$$W_n = \sum_{i=1}^n U(x_i) + \frac{1}{2} \sum_{i \neq j}^n \Phi(x_i, x_j),$$

in which the pair interaction potential is symmetric, $\Phi(x_i, x_j) = \Phi(x_j, x_i)$. The operator of the total energy is the quantum-mechanical *Hamiltonian* $H_n = K_n + W_n$.

In the *field representation*, the general form of an operator acting on the Fock space is

$$\hat{A} = \sum_{n=0}^{\infty} \frac{1}{n!} \int \psi^\dagger(x_n) \dots \psi^\dagger(x_1) A_n(x_1, \dots, x_n) \psi(x_1) \dots \psi(x_n) dx_1 \dots dx_n,$$

with the kernel $A_n(x_1, \dots, x_n)$ being a quantum-mechanical operator. The operators of observables are self-adjoint, $\hat{A}^\dagger = \hat{A}$, since $A_n^\dagger = A_n$.

Examples of the operators of observables in the field representation are the number-of-particle operator

$$\hat{N} = \int \psi^\dagger(x) \psi(x) dx,$$

the momentum operator

$$\hat{\mathbf{P}} = \int \psi^\dagger(x) (-i \nabla) \psi(x) dx,$$

the kinetic-energy operator

$$\hat{K} = \int \psi^\dagger(x) \left(-\frac{\nabla^2}{2m} \right) \psi(x) dx,$$

and the potential-energy operator

$$\hat{W} = \int \psi^\dagger(x) U(x) \psi(x) dx + \frac{1}{2} \int \psi^\dagger(x) \psi^\dagger(x') \Phi(x, x') \psi(x') \psi(x) dx dx'.$$

The total-energy operator is the *Hamiltonian* $\hat{H} = \hat{K} + \hat{W}$ that reads as

$$\hat{H} = \int \psi^\dagger(x) \left[-\frac{\nabla^2}{2m} + U(x) \right] \psi(x) dx + \frac{1}{2} \int \psi^\dagger(x) \psi^\dagger(x') \Phi(x, x') \psi(x') \psi(x) dx dx' .$$

The operators of observables act on the Fock space as

$$\begin{aligned} \hat{N}\varphi &= [nf_n] , \quad \hat{\mathbf{P}}\varphi = [\mathbf{P}_n f_n] , \\ \hat{K}\varphi &= [K_n f_n] , \quad \hat{W}\varphi = [W_n f_n] , \quad \hat{H}\varphi = [H_n f_n] . \end{aligned}$$

In the resulting vector columns, one has quantum-mechanical expressions for $n = 0, 1, 2, \dots$

3.4 Statistical operator

The system statistical state is characterized by a statistical operator $\hat{\rho}(t)$, which is a function of time $t \geq 0$. This is a self-adjoint operator, $\hat{\rho}^+(t) = \hat{\rho}(t)$. It is semipositive and bounded, in the sense of the inequalities

$$0 \leq \frac{\varphi^+ \hat{\rho}(t) \varphi}{\|\varphi\|^2} < \infty$$

for all $\varphi \in \mathcal{F}$ and $\varphi \neq 0$. It is normalized to one, $\text{Tr} \hat{\rho}(t) = 1$, with the trace over the Fock space.

The trace can be taken over any basis $\{\varphi_k\}$ in \mathcal{F} , which is orthonormal, $\varphi_k^+ \varphi_p = \delta_{kp}$, and complete, $\sum_k \varphi_k \varphi_k^+ = \hat{1}$, where $\hat{1}$ is the identity operator in \mathcal{F} . The trace of an operator \hat{A} is given by $\text{Tr} \hat{A} \equiv \sum_k \varphi_k^+ \hat{A} \varphi_k$.

Observable quantities are defined as the *statistical averages*

$$\langle \hat{A}(t) \rangle \equiv \text{Tr} \hat{\rho}(t) \hat{A} . \quad (41)$$

One considers finite systems, with a finite support in real space. Therefore, the operators of observables \hat{A} are termed the operators of local observables. The *algebra of local observables* $\mathcal{A} \equiv \{\hat{A}\}$ is the family of all operators of local observables describing the treated system. And one calls the *statistical state* the set of all observable quantities $\langle \mathcal{A} \rangle \equiv \{ \langle \hat{A} \rangle \}$.

Being defined for a finite system, with \mathbf{r} in volume V , the operator \hat{A} could be written as

$$\hat{A}_V = \sum_{n=0}^{\infty} \frac{1}{n!} \int_V \psi^\dagger(\mathbf{r}_n) \dots \psi^\dagger(\mathbf{r}_1) A_n(\mathbf{r}_1, \dots, \mathbf{r}_n) \psi(\mathbf{r}_1) \dots \psi(\mathbf{r}_n) d\mathbf{r}_1 \dots d\mathbf{r}_n ,$$

where the integration is over the finite system volume. Respectively, the observable quantities are $\langle \hat{A} \rangle_V$. The average number of particles in the system is $N \equiv \langle \hat{N} \rangle_V$.

When taking the thermodynamic limit

$$N \rightarrow \infty , \quad V \rightarrow \infty , \quad \frac{N}{V} \rightarrow \text{const}$$

for extensive quantities that are proportional to N , one should consider the ratio

$$\left| \lim_{N \rightarrow \infty} \frac{1}{N} \langle \hat{A} \rangle_V \right| < \infty .$$

In what follows, the index V in the integrals and averages is omitted, though always keeping in mind that the integration is over the system volume V .

One tells that two operators

$$\hat{A} = \int \hat{A}(\mathbf{r}) d\mathbf{r} , \quad \hat{B} = \int \hat{B}(\mathbf{r}) d\mathbf{r} ,$$

with the densities $\hat{A}(\mathbf{r})$ and $\hat{B}(\mathbf{r})$, satisfy the *clustering property* if

$$\langle \hat{A}(\mathbf{r}) \hat{B}(\mathbf{r}') \rangle \simeq \langle \hat{A}(\mathbf{r}) \rangle \langle \hat{B}(\mathbf{r}') \rangle \quad (42)$$

for $|\mathbf{r} - \mathbf{r}'| \rightarrow \infty$. This property is assumed to hold for the densities of any two operators of observables.

3.5 Evolution operator

In the Schrödinger picture of quantum mechanics, n particles are described by the wave function $\Psi_n(t) = \Psi_n(x_1, \dots, x_n, t)$, where $t \geq 0$ is time. Particles are termed *identical*, provided that

$$|\hat{P}_{ij}\Psi_n(t)|^2 = |\Psi_n(t)|^2 ,$$

where \hat{P}_{ij} is the operator of particle exchange. The *symmetry principle*

$$\hat{P}_{ij}\Psi_n(t) = \pm \Psi_n(t)$$

is a sufficient condition for the particles to be identical. The upper sign corresponds to the Bose-Einstein and the lower, to the Fermi-Dirac statistics.

The scalar product is given by

$$(\Psi_n, \Psi'_n) \equiv \int \Psi_n^*(x_1, \dots, x_n, t) \Psi'_n(x_1, \dots, x_n, t) dx_1 \dots dx_n ,$$

with the norm generated by this scalar product, $\|\Psi_n(t)\| = \sqrt{(\Psi_n, \Psi_n)}$.

The number of particles n is conserved in time, which imposes the condition $\|\Psi_n(t)\| = \|\Psi_n(0)\|$. The wave function pertains to a Hilbert space \mathcal{H}_n and satisfies the Schrödinger equation

$$i \frac{\partial}{\partial t} \Psi_n(t) = H_n \Psi_n(t) .$$

In the Fock space the wave function can be represented as a collection of the Schrödinger wave functions

$$\Psi(t) = \begin{bmatrix} \Psi_0(t) \\ \Psi_1(t) \\ \vdots \\ \Psi_n(t) \\ \vdots \\ \vdots \end{bmatrix} = [\Psi_n(t)] ,$$

with $||\Psi(t)|| = \text{const.}$ The Fock wave function $\Psi(t) \in \mathcal{F}$ satisfies the Schrödinger equation

$$i \frac{d}{dt} \Psi(t) = \hat{H} \Psi(t) ,$$

where

$$\hat{H} \Psi(t) = [H_n \Psi_n(t)] = \left[i \frac{\partial}{\partial t} \Psi_n(t) \right] ,$$

which is equivalent to the set of the Schrödinger equations in \mathcal{H}_n .

The time evolution of the wave function in the Fock space can be represented by means of an *evolution operator* $\hat{U}(t)$ defined so that

$$\Psi(t) = \hat{U}(t) \Psi(0) , \quad \Psi^+(t) = \Psi^+(0) \hat{U}^+(t) .$$

The evolution operator is additive, $\hat{U}(t+t') = \hat{U}(t) \hat{U}(t')$. There exists the identity $\hat{U}(0) = \hat{1}$ and the inverse $\hat{U}^{-1}(t)$. From the norm conservation $||\Psi(t)||^2 = ||\Psi(0)||^2$, we have

$$\Psi^+(t) \Psi(t) = \Psi^+(0) \Psi(0) = \Psi^+(0) \hat{U}^+(t) \hat{U}(t) \Psi(0) ,$$

which shows that $\hat{U}^+(t) \hat{U}(t) = \hat{1}$. Hence the evolution operator is a unitary operator, such that $\hat{U}^+(t) = \hat{U}^{-1}(t)$, with $\hat{U}^{-1}(t) \hat{U}(t) = \hat{1}$. By these properties, the evolution operators, labeled by the time index, form the *evolution group* $\{\hat{U}(t)\}$.

The Schrödinger equation for the wave function $\Psi(t)$ is equivalent to the Schrödinger equation for $\hat{U}(t)$,

$$i \frac{d}{dt} \hat{U}(t) = \hat{H} \hat{U}(t) , \quad i \frac{d}{dt} \hat{U}^+(t) = -\hat{U}^+(t) \hat{H} , \quad (43)$$

where the fact that the Hamiltonian is self-adjoint, $\hat{H}^+ = \hat{H}$, is used.

In general, the evolution operator cannot be represented in explicit form as a function of the Hamiltonian. This is possible, only when the *Lyappo-Danilevsky condition* is valid:

$$\left[\hat{H}, \int_0^t \hat{H} dt \right] = 0 . \quad (44)$$

If this is so, then the formal solution for the evolution operator is

$$\hat{U}(t) = \exp \left\{ -i \int_0^t \hat{H} dt \right\} .$$

This means that the commutator

$$[\hat{H}, \hat{U}(t)] = 0$$

has to be true.

In particular, for \hat{H} not explicitly depending on t ,

$$\hat{U}(t) = \exp \left(-i \hat{H} t \right) .$$

But, in general, the evolution operator cannot be represented in such explicit forms.

3.6 Liouville equation

An n -particle wave function in the Fock space can be written as

$$\varphi_n(t) = \begin{bmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ \Psi_n(t) \\ 0 \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} .$$

A pure-state statistical operator is given by the product $\hat{\rho}_n(t) = \varphi_n(t)\varphi_n^\dagger(t)$. It is the normalized operator, with $\text{Tr}\hat{\rho}_n(t) = 1$, which follows from the wave-function normalizations $\|\Psi_n(t)\| = 1$ and $\|\varphi_n(t)\| = 1$. The pure-state operator is idempotent, $\hat{\rho}_n^2(t) = \hat{\rho}_n(t)$.

A mixed-state statistical operator can be written as

$$\hat{\rho}(t) = \sum_n w_n \varphi_n(t) \varphi_n^\dagger(t) .$$

Since the statistical operator has to be semipositive and normalized, $\text{Tr}\hat{\rho}(t) = 1$, it follows that the weights w_n are to be such that

$$0 \leq w_n \leq 1 , \quad \sum_n w_n = 1 .$$

The time evolution reads as

$$\hat{\rho}(t) = \hat{U}(t) \sum_n w_n \varphi_n(0) \varphi_n^\dagger(0) \hat{U}^\dagger(t) .$$

Denoting the initial statistical operator

$$\hat{\rho}(0) = \sum_n w_n \varphi_n(0) \varphi_n^\dagger(0) ,$$

we come to the form

$$\hat{\rho}(t) = \hat{U}(t) \hat{\rho}(0) \hat{U}^\dagger(t)$$

defining the time evolution of the statistical operator. The latter form leads to the *Liouville equation*

$$i \frac{d}{dt} \hat{\rho}(t) = [\hat{H}, \hat{\rho}(t)] , \tag{45}$$

in which $[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A}$.

3.7 Heisenberg picture

The observable quantities are given by the averages $\langle \hat{A}(t) \rangle \equiv \text{Tr} \hat{\rho}(t) \hat{A}$, which, using the evolution operator, can be written as

$$\text{Tr} \hat{\rho}(t) \hat{A} = \text{Tr} \hat{\rho}(0) \hat{U}^\dagger(t) \hat{A} \hat{U}(t) .$$

The *Heisenberg representation* for the operators of observables is

$$\hat{A}(t) \equiv \hat{U}^\dagger(t) \hat{A}(0) \hat{U}(t) , \quad (46)$$

with the initial condition $\hat{A}(0) = \hat{A}$. Therefore, we have

$$\text{Tr} \hat{\rho}(0) \hat{U}^\dagger(t) \hat{A} \hat{U}(t) = \text{Tr} \hat{\rho}(0) \hat{A}(t) .$$

Then the observable quantities can be written in two equivalent forms:

$$\langle \hat{A}(t) \rangle = \text{Tr} \hat{\rho}(t) \hat{A} = \text{Tr} \hat{\rho}(0) \hat{A}(t) . \quad (47)$$

The relation (46) can also be written as the *Heisenberg equation of motion*

$$i \frac{d}{dt} \hat{A}(t) = \left[\hat{A}(t), \hat{H}(t) \right] . \quad (48)$$

We should notice the difference of the Hamiltonian form in the Heisenberg equation from that in the Liouville equation. In the latter, we have \hat{H} , while the former contains

$$\hat{H}(t) \equiv \hat{U}^\dagger(t) \hat{H} \hat{U}(t) .$$

Generally, $\hat{H}(t)$ does not equal \hat{H} . This happens only when \hat{H} commutes with the evolution operator \hat{U} , which holds true when the Hamiltonian \hat{H} satisfies the Lax-Danilevsky condition (44).

The Heisenberg equation (48) leads to the time evolution of the observable quantities

$$i \frac{d}{dt} \langle \hat{A}(t) \rangle = \langle \left[\hat{A}(t), \hat{H}(t) \right] \rangle .$$

In this way, one can use two representations. In the Schrödinger picture, one deals with $\Psi(t)$, $\hat{\rho}(t)$, and $\hat{A}(0)$. And in the Heisenberg picture, one employs $\Psi(0)$, $\hat{\rho}(0)$, and $\hat{A}(t)$.

The generic form of the operators of observables is

$$\hat{A} = \sum_{n=0}^{\infty} \frac{1}{n!} \int \psi^\dagger(x_n) \dots \psi^\dagger(x_1) A_n(x_1, \dots, x_n) \psi(x_1) \dots \psi(x_n) dx_1 \dots dx_n .$$

The time evolution of the field operator reads as

$$\psi(x, t) \equiv \hat{U}^\dagger(t) \psi(x, 0) \hat{U}(t) , \quad (49)$$

with the initial condition $\psi(x, 0) = \psi(x)$. This is equivalent to the equation of motion

$$i \frac{\partial}{\partial t} \psi(x, t) = \left[\psi(x, t), \hat{H}(t) \right] , \quad (50)$$

where $\hat{H}(t) = \hat{U}^+(t)\hat{H}\hat{U}(t)$.

For the Heisenberg representation of an operator of observable $\hat{A}(t) = \hat{U}^+(t)\hat{A}\hat{U}(t)$, we find

$$\hat{A}(t) = \sum_{n=0}^{\infty} \frac{1}{n!} \int \psi^\dagger(x_n, t) \dots \psi^\dagger(x_1, t) A_n(x_1, \dots, x_n) \psi(x_1, t) \dots \psi(x_n, t) dx_1 \dots dx_n .$$

If the kernel $A_n(x_1, \dots, x_n, t)$ explicitly depends on time, then the Heisenberg equation of motion reads as

$$i \frac{d}{dt} \hat{A}(t) = [\hat{A}(t), \hat{H}(t)] + \hat{U}^+(t) i \frac{\partial \hat{A}}{\partial t} \hat{U}(t) . \quad (51)$$

Again, one has to pay attention that, contrary to the Liouville equation, here we have $\hat{H}(t)$.

The commutation relations for the field operators from Sec. 2.2, can be represented for the time-dependent field operators (49). To this end, we see that

$$[\psi(x, t), \psi(x', t)]_{\mp} = \hat{U}^+(t)[\psi(x), \psi(x')]_{\mp} \hat{U}(t) ,$$

and similarly for other relations. Therefore, we obtain the equal-time commutation relations

$$[\psi(x, t), \psi(x', t)]_{\mp} = 0 , \quad [\psi^\dagger(x, t), \psi^\dagger(x', t)]_{\mp} = 0 , \quad [\psi(x, t), \psi^\dagger(x', t)]_{\mp} = \delta(x - x') . \quad (52)$$

According to Sec. 2.3, the Hamiltonian reads as

$$\begin{aligned} \hat{H}(t) = & \int \psi^\dagger(x, t) \left[-\frac{\nabla^2}{2m} + U(x, t) \right] \psi(x, t) dx + \\ & + \frac{1}{2} \int \psi^\dagger(x, t) \psi^\dagger(x', t) \Phi(x, x') \psi(x', t) \psi(x, t) dx dx' , \end{aligned} \quad (53)$$

so that the equation of motion (50) becomes

$$i \frac{\partial}{\partial t} \psi(x, t) = \left[-\frac{\nabla^2}{2m} + U(x, t) \right] \psi(x, t) + \int \Phi(x, x') \psi^\dagger(x', t) \psi(x', t) \psi(x, t) dx . \quad (54)$$

It is important to stress that the Heisenberg equation of motion (50) for the field operator is equivalent to the variational equation

$$i \frac{\partial}{\partial t} \psi(x, t) = \frac{\delta \hat{H}(t)}{\delta \psi^\dagger(x, t)} . \quad (55)$$

The proof of their equivalence is given in Ref. [43].

3.8 Local dynamics

Using the evolution equations in the Heisenberg representation, it is straightforward to derive the equations for different local density operators [44]. Let us consider the local density of particles

$$\hat{n}(x, t) \equiv \psi^\dagger(x, t) \psi(x, t) . \quad (56)$$

The density commutes with itself, $[\hat{n}(x, t), \hat{n}(x', t)] = 0$, for any statistics. The integral of the density gives the operator of the total number of particles

$$\hat{N} = \int \hat{n}(x, t) dx .$$

And let us define the local current density of particles $\hat{\mathbf{j}} = \hat{\mathbf{j}}(x, t)$ by the equation

$$\hat{\mathbf{j}} \equiv -\frac{i}{2m} [\psi^\dagger(\nabla\psi) - (\nabla\psi^\dagger)\psi] , \quad (57)$$

where, for short, we write $\psi = \psi(x, t)$ and $x = \{\mathbf{r}, \dots\}$.

First, it is easy to derive the *continuity equation*

$$\frac{\partial \hat{n}}{\partial t} + \nabla \cdot \hat{\mathbf{j}} = 0 . \quad (58)$$

The equation for the current density follows:

$$\begin{aligned} \frac{\partial \hat{\mathbf{j}}}{\partial t} + \frac{1}{4m^2} \{ (\nabla^2 \psi^\dagger)(\nabla\psi) - \psi^\dagger \nabla(\nabla^2 \psi) + (\nabla\psi^\dagger)(\nabla^2 \psi) - [\nabla(\nabla^2 \psi^\dagger)] \psi \} = \\ = -\frac{1}{m} \psi^\dagger [\nabla(U + \hat{\Phi})] \psi , \end{aligned} \quad (59)$$

where

$$\hat{\Phi} = \hat{\Phi}(x, t) \equiv \int \Phi(x, x') \hat{n}(x', t) dx' .$$

Analogously, we can get the equation for the time variation of the Hamiltonian density (density of energy)

$$\hat{H}(x, t) \equiv \frac{1}{2m} \nabla\psi^\dagger(x, t) \cdot \nabla\psi(x, t) + \psi(x, t) \left[U(x) + \frac{1}{2} \hat{\Phi}(x, t) \right] \psi(x, t) , \quad (60)$$

which can also be written in the form

$$\hat{H}(x, t) \equiv \psi^\dagger(x, t) \left[-\frac{\nabla^2}{2m} + U(x) + \frac{1}{2} \hat{\Phi}(x, t) \right] \psi(x, t) . \quad (61)$$

The total energy Hamiltonian is $\hat{H} = \int \hat{H}(x, t) dx$.

It is worth emphasizing that the variation in space bears some analogy with the variation in time. To show this, let us treat x as \mathbf{r} . Consider the operator of momentum

$$\hat{\mathbf{P}} = m \int \hat{\mathbf{j}}(\mathbf{r}) d\mathbf{r} = \int \psi^\dagger(\mathbf{r})(-i\nabla)\psi(\mathbf{r}) d\mathbf{r} .$$

We see that the commutator

$$[\psi(\mathbf{r}), \hat{\mathbf{P}}] = -i\nabla\psi(\mathbf{r})$$

reminds us the Heisenberg evolution equation (50). Introducing the *translation operator*

$$\hat{T}(\mathbf{r}) \equiv \exp(i\mathbf{r} \cdot \hat{\mathbf{P}}) , \quad (62)$$

we find

$$\psi(\mathbf{r}) = \hat{T}^+(\mathbf{r})\psi(0)\hat{T}(\mathbf{r}) .$$

This is similar to the dependence of the field operator on time, as expressed through the evolution operator in Eq. (49). That is, the translation operator (62) describes the spatial variation in a manner similar to the description of the time evolution by means of the evolution operator $\hat{U}(t)$.

The translation operator (62) enjoys the properties

$$\hat{T}(0) = \hat{1} , \quad \hat{T}^+(\mathbf{r})\hat{T}(\mathbf{r}) = \hat{1} , \quad \hat{T}^{-1}(\mathbf{r}) = \hat{T}(-\mathbf{r}) = \exp(-i\mathbf{r} \cdot \hat{\mathbf{P}}) , \quad \hat{T}(\mathbf{r})\hat{T}(\mathbf{r}') = \hat{T}(\mathbf{r} + \mathbf{r}') .$$

Hence the family $\{\hat{T}(\mathbf{r})\}$ forms a unitary translation group.

3.9 Coherent states

Coherent states are the eigenstates of the annihilation operator. A coherent state $|\eta\rangle \in \mathcal{F}$ is a solution of the eigenproblem

$$\psi(x)|\eta\rangle = \eta(x)|\eta\rangle , \tag{63}$$

where $\psi(x) \equiv \psi(x, 0)$ and $\eta(x)$ is called *coherent field*. Being a vector of the Fock space, the coherent state has the form

$$|\eta\rangle = \begin{bmatrix} \eta_0 \\ \eta_1(x_1) \\ \eta_2(x_1, x_2) \\ \vdots \\ \eta_n(x_1, \dots, x_n) \\ \vdots \\ \vdots \end{bmatrix} .$$

For short, we shall denote this coherent state as

$$|\eta\rangle = [\eta_n(x_1, \dots, x_n)] .$$

A conjugate to $|\eta\rangle$ is $|\eta\rangle^+ \equiv \langle \eta|$. The scalar product of two coherent states is

$$\langle \eta|\xi\rangle = \sum_{n=0}^{\infty} (\eta_n, \xi_n) ,$$

where the scalar product of the components is

$$(\eta_n, \xi_n) = \int \eta_n^*(x_1, \dots, x_n) \xi_n(x_1, \dots, x_n) dx_1 \dots dx_n .$$

By the definition of the annihilation operator, we have

$$\psi(x)|\eta\rangle = \left[\sqrt{n+1} \eta_{n+1}(x_1, \dots, x_n, x) \right] .$$

Taking into account the definition of the coherent state (63), we get the recursion relation

$$\sqrt{n+1} \eta_{n+1}(x_1, \dots, x_n, x) = \eta(x) \eta_n(x_1, \dots, x_n) .$$

The iterative solution of the latter recursion relation results in the form

$$\eta_n(x_1, \dots, x_n) = \frac{\eta_0}{\sqrt{n!}} \prod_{i=1}^n \eta(x_i) . \quad (64)$$

The scalar product of the components (64) gives

$$(\eta_n, \xi_n) = \frac{\eta_0^* \xi_0}{n!} (\eta, \xi)^n .$$

The scalar product of two coherent states reads as

$$\langle \eta | \xi \rangle = \sum_{n=0}^{\infty} \frac{\eta_0^* \xi_0}{n!} (\eta, \xi)^n = \eta_0^* \xi_0 e^{(\eta, \xi)} ,$$

where

$$(\eta, \xi) \equiv \int \eta^*(x) \xi(x) dx .$$

The coherent state can be normalized as

$$|| |\eta\rangle ||^2 = \langle \eta | \eta \rangle = 1 ,$$

which defines

$$|\eta_0| = \exp \left\{ -\frac{1}{2} (\eta, \eta) \right\} .$$

In this way, the scalar product of two normalized coherent states becomes

$$\langle \eta | \xi \rangle = \exp \left\{ -\frac{1}{2} (\eta, \eta) + (\eta, \xi) - \frac{1}{2} (\xi, \xi) \right\} .$$

Since $0 < |\langle \eta | \xi \rangle| \leq 1$, it follows that the coherent states are not orthogonal with each other. The set $\{|\eta\rangle\}$ forms an overcomplete basis for which the resolution of unity holds true:

$$\int |\eta\rangle \langle \eta| \mathcal{D}\eta = \hat{1} .$$

Integration over the coherent fields can be introduced as follows. Let us take an orthonormal complete basis $\{\chi_k(x)\}$ and let us expand the coherent field as

$$\eta(x) = \sum_k c_k \chi_k(x) , \quad c_k \equiv |c_k| e^{i\alpha_k} ,$$

with α_k real-valued and $0 \leq |c_k| \leq 1$. Then the integration over the complex-valued coherent field is defined with the help of the differential measure

$$\mathcal{D}\eta \equiv \prod_k d|c_k|^2 \frac{d\alpha_k}{2\pi} .$$

A coherent state $|\eta\rangle$ is an eigenstate of the annihilation operator, but it is not an eigenstate of the creation operator $\psi^\dagger(x)$. This is seen from the action

$$\psi^\dagger(x)|\eta\rangle = \left[\sqrt{n} \hat{S}\eta_{n-1}(x_1, \dots, x_{n-1}) \delta(x_n - x) \right] .$$

However, we have the equation

$$\langle \eta | \psi^\dagger(x) = \langle \eta | \eta^*(x) .$$

In view of Eq. (64), the coherent state can be represented as

$$|\eta\rangle = \sum_{n=0}^{\infty} \frac{\eta_0}{n!} \left(\int \eta(x) \psi^\dagger(x) dx \right)^n |0\rangle .$$

In turn, the latter takes the form

$$|\eta\rangle = \eta_0 \exp \left\{ \int \eta(x) \psi^\dagger(x) dx \right\} |0\rangle . \quad (65)$$

Since the coherent state pertains to the ideal Fock space, its components possess the symmetry property

$$\hat{P}_{ij}\eta_n(x_1, \dots, x_n) = \pm \eta_n(x_1, \dots, x_n)$$

with respect to the permutation of the variables x_i and x_j , for $i \neq j$. Therefore the coherent fields satisfy the commutation relations

$$\eta(x_i)\eta(x_j) - \eta(x_j)\eta(x_i) = 0 \quad (\text{Bose})$$

in the case of bosons and the anticommutation relations

$$\eta(x_i)\eta(x_j) + \eta(x_j)\eta(x_i) = 0 \quad (\text{Fermi})$$

in the case of fermions. For bosons, the coherent field is a complex-valued function. While the fermionic coherent fields are Grassmann variables.

Time-dependent coherent states can be defined in the same way as it is done for other states in the Fock space:

$$|\eta(t)\rangle = \hat{U}(t)|\eta\rangle = [\eta_n(x_1, \dots, x_n, t)] . \quad (66)$$

The invariance of the normalization is assumed, so that

$$\langle \eta(t) | \eta(t) \rangle = \langle \eta | \eta \rangle = 1 .$$

By treating the coherent state as an eigenstate of the annihilation operator satisfying the eigenproblem

$$\psi(x)|\eta(t) > = \eta(x, t)|\eta(t) > , \quad (67)$$

we get the definition of a time-dependent coherent field $\eta(x, t)$. The latter equation is equivalent to

$$\psi(x)\hat{U}(t)|\eta > = \eta(x, t)\hat{U}(t)|\eta > ,$$

from where

$$\hat{U}^+(t)\psi(x)\hat{U}(t)|\eta > = \eta(x, t)|\eta > .$$

With $\psi(x, t) = \hat{U}^+(t)\psi(x)\hat{U}(t)$, we obtain another form of Eq. (67) as

$$\psi(x, t)|\eta > = \eta(x, t)|\eta > . \quad (68)$$

Using the time-dependent coherent states, it is possible to pass to the coherent-state representation for the operators of observables. For instance, the number-of-particle operator yields

$$< \eta | \hat{N}(t) | \eta > = \int |\eta(x, t)|^2 dx .$$

And for the Hamiltonian, we find

$$\begin{aligned} < \eta | \hat{H}(t) | \eta > = \int \eta^*(x, t) \left[-\frac{\nabla^2}{2m} + U(x) \right] \eta(x, t) dx + \\ + \frac{1}{2} \int |\eta(x, t)|^2 \Phi(x, x') |\eta(x', t)|^2 dx dx' . \end{aligned}$$

The temporal derivative gives

$$< \eta | i \frac{\partial}{\partial t} \psi(x, t) | \eta > = i \frac{\partial}{\partial t} \eta(x, t) .$$

The evolution equation (55) transforms into the equation

$$i \frac{\partial}{\partial t} \eta(x, t) = \frac{\delta < \eta | \hat{H}(t) | \eta >}{\delta \eta^*(x, t)} . \quad (69)$$

Explicitly, we obtain the *coherent-field equation*

$$i \frac{\partial}{\partial t} \eta(x, t) = \left[-\frac{\nabla^2}{2m} + U(x) + \int \Phi(x, x') |\eta(x', t)|^2 dx' \right] \eta(x, t) . \quad (70)$$

This is the nonlinear Schrödinger equation.

Presenting the solution in the form

$$\eta(x, t) = \eta(x) e^{-iEt} ,$$

we get the stationary coherent-field equation

$$\left[-\frac{\nabla^2}{2m} + U(x) + \int \Phi(x, x') |\eta(x')|^2 dx' \right] \eta(x) = E\eta(x) . \quad (71)$$

It is important to keep in mind that, as is seen, the above equation has a meaning, provided that the interaction potential is integrable, such that

$$\left| \int \Phi(x, x') dx' \right| < \infty .$$

The coherent-field is defined by this equation up to an arbitrary global phase, that is, the general solution is $\eta_\alpha(x, t) = e^{i\alpha} \eta(x, t)$, with α being an arbitrary real-valued quantity.

More information on coherent states can be found in the books [45, 46].

3.10 Density matrices

Statistical averages of the products of field operators define the reduced density matrices [25].

An n -order density matrix is defined as

$$\begin{aligned} \rho_n(x_1, \dots, x_n, x'_1, \dots, x'_n, t) &\equiv \text{Tr } \psi(x_1) \dots \psi(x_n) \hat{\rho}(t) \psi^\dagger(x'_n) \dots \psi^\dagger(x'_1) = \\ &= \text{Tr } \psi(x_1, t) \dots \psi(x_n, t) \hat{\rho} \psi^\dagger(x'_n, t) \dots \psi^\dagger(x'_1, t) = \\ &= \langle \psi^\dagger(x'_n, t) \dots \psi^\dagger(x'_1, t) \psi(x_1, t) \dots \psi(x_n, t) \rangle . \end{aligned} \quad (72)$$

They are symmetric with respect to the simultaneous permutation of x_i with x_j and x'_i with x'_j ,

$$\hat{P}_{ij} \rho_n(x_1, \dots, x_n, x'_1, \dots, x'_n, t) = \rho_n(x_1, \dots, x_n, x'_1, \dots, x'_n, t) .$$

Also, they enjoy the following properties:

$$\begin{aligned} \rho_n^*(x_1, \dots, x_n, x'_1, \dots, x'_n, t) &= \rho_n(x'_n, \dots, x'_1, x_n, \dots, x_1, t) , \\ \rho_n(x_1, \dots, x_n, x'_1, \dots, x'_n, t) &= \rho_n(x_n, \dots, x_1, x'_n, \dots, x'_1, t) , \\ \rho_n^*(x_1, \dots, x_n, x'_1, \dots, x'_n, t) &= \rho_n(x'_1, \dots, x'_n, x_1, \dots, x_n, t) . \end{aligned}$$

The matrices are called *reduced* because their partial integration reduces them to lower-order matrices. Suppose we consider the restricted number-conserving space \mathcal{F}_N , for which the number of particles is fixed, so that for $\varphi_N \in \mathcal{F}_N$, one has $\hat{N}\varphi_N = N\varphi_N$. Then, it follows that $\langle \hat{N}\hat{A} \rangle = N \langle \hat{A} \rangle$. For the second-order density matrix, we have

$$\rho_2(x, x'', x', x'', t) = \langle \hat{n}(x'', t) \psi^\dagger(x', t) \psi(x, t) \rangle - \delta(x' - x'') \rho_1(x, x'', t) ,$$

where $\hat{n}(x, t) \equiv \psi^\dagger(x, t) \psi(x, t)$ is the number-of-particle density. Integrating over x'' yields

$$\int \rho_2(x, x'', x', x'', t) dx'' = \langle \hat{N} \psi^\dagger(x', t) \psi(x, t) \rangle - \rho_1(x, x', t) .$$

On the restricted Fock space \mathcal{F}_N ,

$$\langle \hat{N} \psi^\dagger(x', t) \psi(x, t) \rangle = N \rho_1(x, x', t) .$$

Therefore,

$$\int \rho_2(x, x'', x', x'', t) dx'' = (N - 1) \rho_1(x, x', t) .$$

The name of the matrices contains the term *density*, since the diagonal element of the first-order matrix

$$\rho_1(x, x', t) = \langle \psi^\dagger(x', t) \psi(x, t) \rangle$$

defines the particle density

$$\rho_1(x, x, t) = \langle \psi^\dagger(x, t) \psi(x, t) \rangle \equiv \rho(x, t) .$$

The name *matrices* means that these objects can be treated as matrices. Thus, the matrix

$$\hat{\rho}_n = [\rho_n(x_1, \dots, x_n, x'_1, \dots, x'_n, t)]$$

is a matrix with respect to the x -variables. This matrix is evidently self-adjoint, $\hat{\rho}_n^+ = \hat{\rho}_n$, and semipositive, since

$$\frac{(f_n, \hat{\rho}_n f_n)}{\|f_n\|^2} \geq 0 .$$

Here, the scalar product

$$(f_n, \hat{\rho}_n f_n) \equiv \int f_n^*(x_1, \dots, x_n) \rho(x_1, \dots, x_n, x'_1, \dots, x'_n, t) f_n(x'_1, \dots, x'_n) dx_1 \dots dx_n dx'_1 \dots dx'_n$$

and the vector norm

$$\|f_n\|^2 = (f_n, f_n) = \int f_n^*(x_1, \dots, x_n) f_n(x_1, \dots, x_n) dx_1 \dots dx_n$$

are used.

The norm of an n -order density matrix is given by the expression

$$\|\hat{\rho}_n\| \equiv \sup_{\|f_n\|=1} |(f_n, \hat{\rho}_n f_n)| .$$

The eigenproblem

$$\hat{\rho}_n f_{nj} = \lambda_{nj} f_{nj} ,$$

implies the equation

$$\int \rho_n(x_1, \dots, x_n, x'_1, \dots, x'_n, t) f_{nj}(x'_1, \dots, x'_n) dx'_1 \dots dx'_n = \lambda_{nj} f_{nj}(x_1, \dots, x_n) .$$

Being self-adjoint and semipositive, the density matrix enjoys real non-negative eigenvalues. The largest of the eigenvalues defining the norm

$$\|\hat{\rho}_n\| = \sup_j \lambda_{nj} \equiv \lambda_n$$

is of special importance characterizing the level of ordering in the system.

The normalization for the density matrices can be summarized as follows:

$$\text{Tr} \hat{\rho}_1 \equiv \int \rho_1(x, x, t) dx = N ,$$

$$\text{Tr} \hat{\rho}_2 \equiv \int \rho_2(x_1, x_2, x_1, x_2, t) dx_1 dx_2 = N(N-1) ,$$

$$\text{Tr} \hat{\rho}_n \equiv \int \rho_n(x_1, \dots, x_n, x_1, \dots, x_n, t) dx_1 \dots dx_n = \frac{N!}{(N-n)!}.$$

The level of ordering in the system can be described [25, 26] by the *order indices*

$$\omega(\hat{\rho}_n) \equiv \frac{\log \|\hat{\rho}_n\|}{\log \text{Tr} \hat{\rho}_n}. \quad (73)$$

Keeping in mind that $N \gg n \geq 1$, we may use the Stirling formula $z! \simeq \sqrt{2\pi} e^{-z} z^{z+1/2}$, in which $|z| \rightarrow \infty$. Then, we have

$$\frac{N!}{(N-n)!} \simeq \left(\frac{N}{e}\right)^n, \quad \log \frac{N!}{(N-n)!} \simeq n \log N.$$

As a result, the order index (73) reads as

$$\omega(\hat{\rho}_n) = \frac{\log \lambda_n}{n \log N}, \quad (74)$$

for $N \gg n \geq 1$.

The order indices allow us to distinguish between long-range order, mid-range order, and short-range order [25, 26]. This is contrary to the order parameters that distinguish only the long-range order from the absence of such an order.

By means of the reduced density matrices, it is possible to express the statistical averages of operators of observables. In the general case, we have

$$\langle \hat{A}(t) \rangle = \sum_{n=0}^{\infty} \frac{1}{n!} \int \lim_{\{x'_i \rightarrow x_i\}} A_n(x_1, \dots, x_n) \rho_n(x_1, \dots, x_n, x'_1, \dots, x'_n, t) dx_1 \dots dx_n.$$

In particular cases, for the average of the number-of-particle operator, we get

$$\langle \hat{N} \rangle = \int \rho_1(x, x, t) dx = N.$$

For the kinetic-energy operator,

$$\langle \hat{K} \rangle = \int \lim_{x' \rightarrow x} \left(-\frac{\nabla^2}{2m} \right) \rho_1(x, x', t) dx.$$

And for the potential-energy operator

$$\langle \hat{W} \rangle = \int U(x) \rho_1(x, x, t) dx + \frac{1}{2} \int \Phi(x_1, x_2) \rho_2(x_1, x_2, x_1, x_2, t) dx_1 dx_2.$$

The total internal energy is $E \equiv \langle \hat{H} \rangle = \langle \hat{K} \rangle + \langle \hat{W} \rangle$.

The equations of motion for the reduced density matrices follow from their definition and Eq. (54) for the field operator. The first-order density matrix satisfies the equation

$$i \frac{\partial}{\partial t} \rho_1(x_1, x_2, t) = \left[-\frac{\nabla_1^2}{2m} + U(x_1) + \frac{\nabla_2^2}{2m} - U(x_2) \right] \rho_1(x_1, x_2, t) + \quad (75)$$

$$+ \int [\Phi(x_1, x_3) - \Phi(x_2, x_3)] \rho_2(x_1, x_3, x_2, x_3, t) dx_3 ,$$

involving the second-order matrix. The equation for the second-order matrix involves the third-order matrix, and so on. The sequence of these equations is called the *Bogolubov - Born - Green - Kirkwood - Yvon hierarchy* [44, 47–49].

Particle correlations are characterized by the *pair correlation function*

$$g(x, x', t) \equiv \frac{\langle \psi^\dagger(x, t) \psi^\dagger(x', t) \psi(x', t) \psi(x, t) \rangle}{\rho(x, t) \rho(x', t)} ,$$

which is expressed through the second-order density matrix:

$$g(x, x', t) = \frac{\rho_2(x, x', x, x', t)}{\rho(x, t) \rho(x', t)} .$$

The pair correlation function is real, $g^*(x, x', t) = g(x, x', t)$ and symmetric, $g(x', x, t) = g(x, x', t)$. It satisfies the asymptotic condition

$$g(x, x', t) \rightarrow 1 \quad (|x - x'| \rightarrow \infty),$$

following from the clustering property (42).

4 Statistical states

In this section, different types of statistical states are treated and their properties are discussed. Several important mathematical facts are presented, which are widely used for describing quantum systems.

4.1 Gibbs entropy

Gibbs entropy is the central notion for characterizing statistical states. It is also called Shannon entropy or information entropy, since it quantifies the amount of the unavailable information on the considered system. For a statistical operator $\hat{\rho}$, the Gibbs entropy is

$$S(t) \equiv -\text{Tr} \hat{\rho}(t) \ln \hat{\rho}(t) . \quad (76)$$

Since $\hat{\rho}(t)$ is semipositive, bounded, and normalized to unity, $S(t)$ is semipositive, $S(t) \geq 0$.

Here and in what follows, the functions of operators are defined by the rule

$$f(\hat{A}) \equiv \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} (\hat{A} - a)^n , \quad f^{(n)}(a) \equiv \frac{d^n f(a)}{da^n} ,$$

where a is a convenient nonoperator quantity. Below, several important facts, related to the Gibbs entropy are proved.

Liouville theorem. Gibbs entropy does not change with time:

$$S(t) = S(0) . \quad (77)$$

Proof. The time evolution of the statistical operator is given by the relation

$$\hat{\rho}(t) = \hat{U}(t) \hat{\rho} \hat{U}^\dagger(t) ,$$

where $\hat{\rho} \equiv \hat{\rho}(0)$. By the definition of the function of an operator, we have

$$\ln \hat{\rho}(t) = \hat{U}(t) \ln \hat{\rho} \hat{U}^\dagger(t) .$$

Therefore

$$S(t) = -\text{Tr} \hat{\rho}(0) \ln \hat{\rho}(0) = S(0) ,$$

which proves equality (77). \square

Gibbs inequality. Let \hat{A} and \hat{B} be semi-positive operators, possessing traces, then

$$\text{Tr} \hat{A} \ln \hat{A} - \text{Tr} \hat{A} \ln \hat{B} \geq \text{Tr} (\hat{A} - \hat{B}) . \quad (78)$$

Proof. For a semi-positive operator \hat{A} , one has

$$\ln \hat{A} \geq 1 - \hat{A}^{-1} .$$

Similarly,

$$\ln \hat{A} - \ln \hat{B} = \ln(\hat{A}\hat{B}^{-1}) \geq 1 - \hat{A}^{-1}\hat{B} .$$

From here,

$$\hat{A} \ln(\hat{A}\hat{B}^{-1}) \geq \hat{A} - \hat{B} .$$

Then inequality (78) follows. \square

Let us introduce the *entropy operator*

$$\hat{S}(t) \equiv -\ln \hat{\rho}(t) . \quad (79)$$

And let us define the *average entropy*

$$\overline{S}(t) \equiv \text{Tr} \hat{\rho}(0) \hat{S}(t) . \quad (80)$$

Notice that the average entropy does not equal $\text{Tr} \hat{\rho}(t) \hat{S}(0)$.

Nondecrease of average entropy. The average entropy (80) does not decrease with time:

$$\overline{S}(t) \geq \overline{S}(0) . \quad (81)$$

Proof. By the definition of the average entropy, one has

$$\overline{S}(t) - \overline{S}(0) = \text{Tr} \hat{\rho} [\ln \hat{\rho} - \ln \hat{\rho}(t)] .$$

The proof follows the same arguments as in the proof of the Gibbs inequality. We have

$$\ln \hat{\rho} - \ln \hat{\rho}(t) = \ln \hat{\rho} \hat{\rho}^{-1}(t) .$$

Since

$$\ln \hat{\rho} \hat{\rho}^{-1}(t) \geq 1 - \hat{\rho}^{-1} \hat{\rho}(t) ,$$

we get

$$\overline{S}(t) - \overline{S}(0) \geq \text{Tr} \hat{\rho} [1 - \hat{\rho}^{-1} \hat{\rho}(t)] .$$

Then

$$\text{Tr} \hat{\rho} [1 - \hat{\rho}^{-1} \hat{\rho}(t)] = \text{Tr} \hat{\rho} - \text{Tr} \hat{\rho}(t) .$$

As far as $\text{Tr} \hat{\rho} = \text{Tr} \hat{\rho}(t) = 1$, we come to the inequality

$$\overline{S}(t) - \overline{S}(0) \geq 0 ,$$

which yields Eq. (81). \square

4.2 Information functionals

The Gibbs entropy serves as an efficient tool for defining the form of the statistical operators. This is done employing the *principle of minimal information*, which implies a conditional maximum of entropy under given additional constraints. The idea of this principle goes back to Gibbs [34, 35, 50] and has also been considered by Shannon [51] and Janes [52]. General formulation and different applications have been studied in Refs. [33, 46].

Suppose that some operators of local observables $\hat{A}_i \in \mathcal{A}$ at $t = 0$ are assumed to give the known observable quantities

$$A_i \equiv \langle \hat{A}_i \rangle = \text{Tr} \hat{\rho} \hat{A}_i , \quad (82)$$

enumerated with $i = 0, 1, 2, \dots$. One of the quantities is always $\hat{A}_0 \equiv \hat{1}$, which guarantees the normalization condition

$$\langle \hat{1} \rangle = \text{Tr} \hat{\rho} = 1 .$$

The operators entering constraint (82) are termed the *constraint operators*. The information functional, or simply information, is

$$I[\hat{\rho}] \equiv -S + \sum_i \lambda_i \left(\text{Tr} \hat{\rho} \hat{A}_i - A_i \right) , \quad (83)$$

where λ_i are Lagrange multipliers, controlling the validity of conditions (82) through the variation $\delta I[\hat{\rho}]/\delta \lambda_i = 0$. The statistical operator is supposed to be such that to provide the minimum of information (83). Minimization with respect to $\hat{\rho}$ implies

$$\frac{\delta I[\hat{\rho}]}{\delta \hat{\rho}} = 0 , \quad \frac{\delta^2 I[\hat{\rho}]}{\delta \hat{\rho}^2} > 0 .$$

Using the derivatives

$$\frac{\delta}{\delta \hat{\rho}} \text{Tr} \hat{\rho} \hat{A}_i = \hat{A}_i , \quad \frac{\delta}{\delta \hat{\rho}} \ln \hat{\rho} = \hat{\rho}^{-1}$$

leads to the *Gibbs statistical operator*

$$\hat{\rho} = \frac{1}{Z} \exp \left(- \sum_{i>0} \lambda_i \hat{A}_i \right) , \quad (84)$$

where

$$Z = \text{Tr} \exp \left(- \sum_{i>0} \lambda_i \hat{A}_i \right)$$

is called the *partition function*. The second derivative is positive, since

$$\frac{\delta^2 I[\hat{\rho}]}{\delta \hat{\rho}^2} = \hat{\rho}^{-1} > 0 .$$

In the same way, it is possible to find the statistical operator, when the local operator densities at $t = 0$ yield the known densities of observables

$$A_i(x) \equiv \langle \hat{A}_i(x) \rangle = \text{Tr} \hat{\rho} \hat{A}_i(x) . \quad (85)$$

Then one defines the local information functional

$$I_{loc}[\hat{\rho}] \equiv -S + \sum_i \int \lambda_i(x) \left[\text{Tr} \hat{\rho} \hat{A}_i(x) - A_i(x) \right] dx . \quad (86)$$

The minimization of information (86) results in the statistical operator

$$\hat{\rho} = \frac{1}{Z} \exp \left\{ - \sum_{i>0} \int \lambda_i(x) \hat{A}_i(x) dx \right\} , \quad (87)$$

with the partition function

$$Z = \text{Tr} \exp \left\{ - \sum_{i>0} \int \lambda_i(x) \hat{A}_i(x) dx \right\} .$$

The statistical operator (87) is called the *local-equilibrium* or *quasi-equilibrium* statistical operator.

4.3 Equilibrium states

A *statistical state* is the set $\{\langle \hat{A}(t) \rangle\}$ of the averages of all operators of local observables. By this definition, a statistical state possesses several general properties that are studied in the algebraic approach [53].

An *equilibrium state* does not depend on time, so that $\langle \hat{A}(t) \rangle = \langle \hat{A}(0) \rangle$ for all averages from the state $\{\langle \hat{A}(t) \rangle\}$. This implies that $\hat{\rho}(t) = \hat{\rho}(0) = \hat{\rho}$, because of which $[\hat{H}, \hat{\rho}] = 0$. The latter means that the statistical operator $\hat{\rho}$ is an integral of motion. Hence it has to be a function of \hat{H} and other integrals of motion. An *integral of motion* \hat{C} is an operator for which

$$[\hat{C}, \hat{H}] = 0 , \quad i \frac{d}{dt} \hat{C} = [\hat{C}, \hat{H}] = 0 .$$

The Hamiltonian \hat{H} is, clearly, also an integral of motion.

Replacing in the information functionals \hat{A}_i by \hat{C}_i , and minimizing the information, we get the *Gibbs statistical operator*

$$\hat{\rho} = \frac{1}{Z} \exp \left(- \sum_i \lambda_i \hat{C}_i \right) ,$$

with the partition function

$$Z = \text{Tr} \exp \left(- \sum_i \lambda_i \hat{C}_i \right) .$$

The statistical state $\{ \langle \hat{A} \rangle \}$, composed of the averages $\langle \hat{A} \rangle \equiv \text{Tr} \hat{\rho} \hat{A}$, with the Gibbs statistical operator, is termed the *Gibbs state*.

One says that $\hat{\rho}$ defines a statistical ensemble, which is a collection of possible dynamical states of a physical system. More rigorously, a statistical ensemble is a pair $\{ \mathcal{F}, \hat{\rho} \}$ of the Fock space of microscopic states \mathcal{F} and a statistical operator $\hat{\rho}$. Depending on the choice of the constraint operators, there are different statistical ensembles represented by different statistical operators.

If, in addition to the normalization condition, one considers the sole constraint operator that is the Hamiltonian energy operator \hat{H} , then the principle of minimal information yields the *Gibbs canonical operator*

$$\hat{\rho} = \frac{1}{Z} e^{-\beta \hat{H}} , \quad Z = \text{Tr} e^{-\beta \hat{H}} , \quad (88)$$

where β is inverse temperature, $\beta T \equiv 1$. The statistical operator $\hat{\rho}$ describes the *Gibbs canonical ensemble*. In this ensemble, the role of a thermodynamic potential is played by the *free energy* $F = F(T, V, N, \mathbf{M})$, which can be defined by the equality

$$\text{Tr} e^{-\beta \hat{H}} \equiv e^{-\beta F} .$$

Therefore

$$F \equiv -T \ln Z = -T \ln \text{Tr} e^{-\beta \hat{H}} . \quad (89)$$

Other thermodynamic quantities can be found through the free energy, for instance

$$S = - \left(\frac{\partial F}{\partial T} \right)_{VN} = \beta \left(\langle \hat{H} \rangle - F \right) , \quad E \equiv \langle \hat{H} \rangle = F + TS .$$

The Gibbs entropy (information entropy) coincides with the thermodynamic entropy,

$$S = -\text{Tr} \hat{\rho} \ln \hat{\rho} = \beta(E - F) .$$

When, in addition to the normalization condition, as the constraint operators one takes the Hamiltonian energy \hat{H} and the number-of-particle operator \hat{N} , then one gets the *grand canonical operator*

$$\hat{\rho} = \frac{1}{Z} e^{-\beta(\hat{H} - \mu \hat{N})} , \quad Z = \text{Tr} e^{-\beta(\hat{H} - \mu \hat{N})} , \quad (90)$$

where μ is chemical potential. This statistical operator describes the *grand canonical ensemble*. The corresponding thermodynamic potential is the grand potential $\Omega = \Omega(T, V, \mu, \mathbf{M})$ defined by the equality

$$\text{Tr} e^{-\beta(\hat{H} - \mu \hat{N})} \equiv e^{-\beta \Omega} .$$

This means that

$$\Omega = -T \ln \text{Tr} e^{-\beta(\hat{H} - \mu \hat{N})} . \quad (91)$$

The grand potential $\Omega = F - \mu N$ defines other thermodynamic quantities, for example, the entropy

$$S = - \left(\frac{\partial \Omega}{\partial T} \right)_{V\mu} = \beta(E - \mu N - \Omega) .$$

The Gibbs and thermodynamic entropies coincide, so that the above expression of the entropy equals $S = -\text{Tr} \hat{\rho} \ln \hat{\rho}$.

Generally, different statistical ensembles can describe different physical situations. If the considered systems are physically different, then one of the ensembles may be appropriate for characterizing one of the system, while another ensemble may be not appropriate for the same system, but appropriate for the other system. In that sense, the statistical ensembles do not need to be completely equivalent. Though in the majority of the cases, the ensembles are equivalent, so that any of them can be used.

It is possible to distinguish the ensemble equivalence in two senses. One meaning is the *thermodynamic equivalence of ensembles*, when the thermodynamic characteristics, calculated for different ensembles, coincide, after being expressed through the same variables, so that

$$S = - \left(\frac{\partial F}{\partial T} \right)_{VN} = - \left(\frac{\partial \Omega}{\partial T} \right)_{V\mu} , \quad P = - \left(\frac{\partial F}{\partial V} \right)_{TN} = - \left(\frac{\partial \Omega}{\partial V} \right)_{T\mu} .$$

Another meaning is the *statistical equivalence of ensembles*, when the statistical states coincide, so that the observable quantities calculated in one ensemble, say in the canonical ensemble giving $\langle \hat{A} \rangle_N$, coincide with the same observable calculated in another ensemble, say in the grand canonical ensemble giving $\langle \hat{A} \rangle_\mu$, after taking into account the relation between the chemical potential and the number of particles.

4.4 Representative ensembles

In order to correctly characterize a system, one has to chose a *representative statistical ensemble*, which is an ensemble that uniquely describes the considered system [33,34,54,55]. Actually, the problem of ensemble equivalence is meaningful only for representative ensembles [55].

An *equilibrium statistical ensemble* is a pair $\{\mathcal{F}, \hat{\rho}\}$ of the Fock space and a statistical operator. To uniquely describe the system, the ensemble must be representative, which requires that the statistical operator has to be defined taking into account all constraints that are necessary for the unique system description. For this, one has to consider the necessary constraint operators, $\hat{C}_i^+ = \hat{C}_i$ that are assumed to be self-adjoint, but not necessarily integrals of motion. These constraint operators define the corresponding *statistical conditions*

$$C_i = \langle \hat{C}_i \rangle = \text{Tr} \hat{\rho} \hat{C}_i , \quad (92)$$

with the trace over \mathcal{F} .

The first standard constraint is the normalization condition for the statistical operator,

$$1 = \langle \hat{1}_{\mathcal{F}} \rangle = \text{Tr} \hat{\rho} ,$$

where $\hat{1}_{\mathcal{F}}$ is a unity operator in \mathcal{F} . Another usual condition is the definition of the internal energy

$$E = \langle \hat{H} \rangle = \text{Tr} \hat{\rho} \hat{H} .$$

Keeping in mind the normalization condition, the definition of the internal energy, and other constraints (92), required for the unique description of the system, we construct the information functional

$$I[\hat{\rho}] = \text{Tr} \hat{\rho} \ln \hat{\rho} + \lambda_0 (\text{Tr} \hat{\rho} - 1) + \beta (\text{Tr} \hat{\rho} \hat{H} - E) + \beta \sum_i \nu_i (\text{Tr} \hat{\rho} \hat{C}_i - C_i) ,$$

in which λ_0 , β , and $\beta\nu_i$ are the Lagrange multipliers.

The minimization of the information functional gives the statistical operator

$$\hat{\rho} = \frac{1}{Z} e^{-\beta H} , \quad (93)$$

in which the *grand Hamiltonian*

$$H = \hat{H} + \sum_i \nu_i \hat{C}_i \quad (94)$$

is introduced. The partition function is $Z = \exp(1 + \lambda_0) = \text{Tr} \exp -\beta H$. Such a statistical operator defines the representative statistical ensemble $\{\mathcal{F}, \hat{\rho}\}$.

4.5 Action functional

A nonequilibrium statistical ensemble is a pair $\{\mathcal{F}, \hat{\rho}(t)$, where the statistical operator is a function of time characterized by an evolution operator and an initial condition $\hat{\rho} \equiv \hat{\rho}(0)$.

In quantum field theory, the system evolution is defined by an action functional [56]. The temporal energy operator is written in the form

$$\hat{E} \equiv \int \psi^\dagger(x, t) i \frac{\partial}{\partial t} \psi(x, t) dx . \quad (95)$$

The system dynamical properties are described by the Lagrangian $\hat{L} \equiv \hat{E} - \hat{H}$ and the *action functional*

$$A[\psi] \equiv \int (\hat{L} - \nu_i \hat{C}_i) dt = \int (\hat{E} - H) dt , \quad (96)$$

taking into account the constraints that uniquely define the system. Here H is the grand Hamiltonian (94).

The equations of motion are given by the extremum of the action functional,

$$\frac{\delta A[\psi]}{\delta \psi^\dagger(x, t)} = 0 , \quad \frac{\delta A[\psi]}{\delta \psi(x, t)} = 0 . \quad (97)$$

From the form of the action functional (96), it follows that the extremization of the action functional is identical to the variational equation

$$i \frac{\partial}{\partial t} \psi(x, t) = \frac{\delta H}{\delta \psi^\dagger(x, t)} . \quad (98)$$

And the latter, as has been proved [43], is equivalent to the Heisenberg equation of motion

$$i \frac{\partial}{\partial t} \psi(x, t) = [\psi(x, t), H] . \quad (99)$$

A representative ensemble takes into account all statistical conditions, necessary for the unique description of the considered statistical system. By defining a representative ensemble, one comes to the notion of the grand Hamiltonian. And, as is shown above, the grand Hamiltonian governs both the statistical properties of equilibrium systems as well as the temporal evolution of nonequilibrium systems. This means that the Liouville equation for the statistical operator must also contain the grand Hamiltonian.

4.6 Fluctuations and stability

In section 2.2, it is explained that in equilibrium systems, susceptibilities are to be positive and finite. Here we show how the susceptibilities are connected with fluctuations of observable quantities.

Observable quantities are represented by self-adjoint operators. Fluctuations of an observable, corresponding to an operator \hat{A} , are characterized by the *variance* or *dispersion*

$$\text{var}(\hat{A}) \equiv \langle \hat{A}^+ \hat{A} \rangle - |\langle \hat{A} \rangle|^2 . \quad (100)$$

Generally, variance (100) can be defined for any \hat{A} . For a self-adjoint $\hat{A}^+ = \hat{A}$, it reduces to

$$\text{var}(\hat{A}) \equiv \langle \hat{A}^2 \rangle - \langle \hat{A} \rangle^2 . \quad (101)$$

By this definition, $\text{var}(\hat{A}) \geq 0$, since

$$\text{var}(\hat{A}) = \langle (\hat{A} - \langle \hat{A} \rangle)^2 \rangle .$$

That is why susceptibilities are to be non-negative.

Let us consider *specific heat*

$$C_V = \frac{1}{N} \left(\frac{\partial E}{\partial T} \right)_{VN} = \frac{T}{N} \left(\frac{\partial S}{\partial T} \right)_{VN} ,$$

which can also be expressed through the second derivatives

$$C_V = - \frac{T}{N} \left(\frac{\partial^2 F}{\partial T^2} \right)_{VN} = - \frac{T}{N} \left(\frac{\partial^2 \Omega}{\partial T^2} \right)_{V\mu} .$$

Using this definition, for the specific heat we have

$$C_V = \frac{\text{var}(\hat{H})}{NT^2} , \quad (102)$$

which shows that it characterizes energy fluctuations.

Consider now the *isothermal compressibility*

$$\kappa_T = - \frac{1}{V} \left(\frac{\partial V}{\partial P} \right)_{TN} = \frac{1}{N\rho} \left(\frac{\partial N}{\partial \mu} \right)_{TV} ,$$

which can also be written as

$$\kappa_T = - \frac{1}{N\rho} \left(\frac{\partial^2 \Omega}{\partial \mu^2} \right)_{TV} = \frac{1}{\rho^2} \left(\frac{\partial^2 P}{\partial \mu^2} \right)_{TV} .$$

Therefore for the compressibility, we get

$$\kappa_T = \frac{\text{var}(\hat{N})}{N\rho T} , \quad (103)$$

which means that it describes the number-of-particle fluctuations.

The number-of-particle variance can be written as

$$\text{var}(\hat{N}) = \int [\langle \hat{n}(x)\hat{n}(x') \rangle - \rho(x)\rho(x')] dx dx' ,$$

being expressed through the density-density correlation function

$$\langle \hat{n}(x)\hat{n}(x') \rangle = \rho_2(x, x', x, x') + \rho(x)\delta(x - x') .$$

Therefore, it can be written as

$$\text{var}(\hat{N}) = N + \int \rho(x)\rho(x') [g(x, x') - 1] dx dx' . \quad (104)$$

Considering the *magnetic susceptibility*, we take into account that an external magnetic field \mathbf{B} enters the Hamiltonian through the expression $-\sum_{i=1}^N \hat{\mathbf{M}}_i \cdot \mathbf{B}$, where $\hat{\mathbf{M}}_i$ is the magnetic-moment operator of an i -particle. The magnetic susceptibility tensor is defined as

$$\chi_{\alpha\beta} = \frac{1}{N} \frac{\partial M^\alpha}{\partial B^\beta} ,$$

where the components of the average magnetization are

$$M^\alpha = - \frac{\partial G}{\partial B^\alpha} = \langle \hat{M}^\alpha \rangle , \quad \hat{\mathbf{M}} \equiv \sum_{i=1}^N \hat{\mathbf{M}}_i .$$

Then the susceptibility tensor is

$$\chi_{\alpha\beta} = - \frac{1}{N} \frac{\partial^2 G}{\partial B^\alpha \partial B^\beta} = \frac{\langle \hat{M}^\alpha \hat{M}^\beta \rangle - \langle \hat{M}^\alpha \rangle \langle \hat{M}^\beta \rangle}{NT} .$$

The diagonal elements of the tensor give the magnetic susceptibilities

$$\chi_{\alpha\alpha} = \frac{\text{var}(\hat{M}^\alpha)}{NT} , \quad (105)$$

which shows that they describe magnetic-moment fluctuations.

Being proportional to the variances of the related observables, all these susceptibilities are non-negative:

$$C_V \geq 0 , \quad \kappa_T \geq 0 , \quad \chi_{\alpha\alpha} \geq 0 ,$$

which is a part of the stability conditions. For the stability of an equilibrium system it is also required that the susceptibilities be finite:

$$C_V < \infty , \quad \kappa_T < \infty , \quad \chi_{\alpha\alpha} < \infty .$$

This means that the stability conditions, in terms of the variance of observables, are

$$0 \leq \frac{\text{var}(\hat{A})}{N} < \infty . \quad (106)$$

This has to be valid for any N , including $N \rightarrow \infty$.

When $\text{var}(\hat{A}) \propto N$, this is called the *thermodynamically normal* dispersion corresponding to thermodynamically normal fluctuations. And if $\text{var}(\hat{A}) \propto N^\gamma$, with $\gamma > 1$, this is called the *thermodynamically anomalous* dispersion describing thermodynamically anomalous fluctuations. Thermodynamically anomalous fluctuations imply the system instability.

Summarizing, the susceptibilities

$$C_V = \frac{\text{var}(\hat{H})}{NT^2} , \quad \kappa_T = \frac{\text{var}(\hat{N})}{N\rho T} , \quad \chi_{\alpha\alpha} = \frac{\text{var}(\hat{M}^\alpha)}{NT}$$

describe the fluctuations of the corresponding observables, they are expressed through the variance of the related operators and have to be non-negative and finite for any equilibrium system.

4.7 Gibbs-Bogolubov inequality

A very important result, widely used in statistical physics is the inequality, first, proved by Gibbs for classical systems and then extended by Bogolubov to quantum systems, because of which it is named the Gibbs-Bogolubov inequality.

Let us consider two self-adjoint operators \hat{A} and \hat{B} . And let us define the functionals

$$F(\hat{A}) \equiv -T \ln \text{Tr} e^{-\beta \hat{A}} , \quad F(\hat{B}) \equiv -T \ln \text{Tr} e^{-\beta \hat{B}} .$$

We introduce the averages

$$\langle \hat{A} \rangle_A \equiv \text{Tr} \hat{\rho}_A \hat{A} , \quad \langle \hat{A} \rangle_B \equiv \text{Tr} \hat{\rho}_B \hat{A} ,$$

in which

$$\hat{\rho}_A \equiv \frac{e^{-\beta \hat{A}}}{\text{Tr} e^{-\beta \hat{A}}} , \quad \hat{\rho}_B \equiv \frac{e^{-\beta \hat{B}}}{\text{Tr} e^{-\beta \hat{B}}} .$$

Then the following theorem is valid.

Gibbs-Bogolubov inequality. For the quantities $F(\hat{A})$ and $F(\hat{B})$, one has

$$F(\hat{A}) \leq F(\hat{B}) + \langle \hat{A} - \hat{B} \rangle_B \quad (107)$$

and

$$\langle \hat{A} - \hat{B} \rangle_A + F(\hat{B}) \leq F(\hat{A}) . \quad (108)$$

Proof. According to the above definitions, we have

$$\text{Tr} \hat{\rho}_B \ln \hat{\rho}_B = \beta \left[F(\hat{B}) - \langle \hat{B} \rangle_B \right] ,$$

$$\text{Tr} \hat{\rho}_B \ln \hat{\rho}_A = \beta \left[F(\hat{A}) - \langle \hat{A} \rangle_B \right] .$$

By the Gibbs inequality (78),

$$\text{Tr} \hat{\rho}_B \ln \hat{\rho}_B \geq \text{Tr} \hat{\rho}_B \ln \hat{\rho}_A .$$

From here, we get inequality (107). Interchanging \hat{A} and \hat{B} , we come to inequality (108). \square

The most often used application of the Gibbs-Bogolubov inequality is for finding the best variational approximation for a complicated problem. Suppose we need to calculate the free energy

$$F(\hat{H}) = -T \ln \text{Tr} e^{-\beta \hat{H}} ,$$

which is impossible to calculate exactly. Then we can take an approximate free energy

$$F(\hat{H}_{app}) = -T \ln \text{Tr} e^{-\beta \hat{H}_{app}} ,$$

defined through an approximate Hamiltonian H_{app} . We use the notation

$$\langle \hat{A} \rangle = \text{Tr} \hat{\rho} \hat{A} , \quad \hat{\rho} = \frac{\exp(-\beta \hat{H})}{\text{Tr} \exp(-\beta \hat{H})} ,$$

$$\langle \hat{A} \rangle_{app} = \text{Tr} \hat{\rho}_{app} \hat{A} , \quad \hat{\rho}_{app} = \frac{\exp(-\beta \hat{H}_{app})}{\text{Tr} \exp(-\beta \hat{H}_{app})} .$$

Then, by the Gibbs-Bogolubov inequality,

$$F(\hat{H}) \leq F(\hat{H}_{app}) + \langle \hat{H} - \hat{H}_{app} \rangle_{app} . \quad (109)$$

The best approximation is obtained by minimizing the right-hand side of inequality (109), that is, by choosing the approximate Hamiltonian that provides such a minimization.

A particular form of the Gibbs-Bogolubov inequality is derived for zero temperature, when $T \rightarrow 0$, hence $\beta \rightarrow \infty$. By the definition of the trace,

$$\text{Tr} e^{-\beta \hat{H}} = \sum_k \varphi_k^+ e^{-\beta \hat{H}} \varphi_k ,$$

where φ_k can be taken as the eigenfunctions of the Hamiltonian,

$$\hat{H} \varphi_k = E_k \varphi_k .$$

The *ground-state energy* is

$$E_0 \equiv \min_k E_k = (\varphi_0, \hat{H} \varphi_0) .$$

Assuming a non-degenerate spectrum, the above trace can be rewritten as

$$\sum_k \varphi_k^+ e^{-\beta \hat{H}} \varphi_k = e^{-\beta E_0} \left[1 + \sum_{k \neq 0} e^{-\beta(E_k - E_0)} \right] .$$

Since $E_k - E_0 > 0$ for $k \neq 0$, we have

$$\text{Tr} e^{-\beta \hat{H}} \simeq e^{-\beta E_0} \quad (\beta \rightarrow \infty) .$$

Therefore

$$F(\hat{H}) \simeq E_0 \quad (T \rightarrow 0) .$$

Respectively, when temperature tends to zero, we get

$$\langle \hat{A} \rangle \simeq (\varphi_0, \hat{A}\varphi_0) \quad (T \rightarrow 0) .$$

For the approximate Hamiltonian, one has

$$\hat{H}_{app}\varphi_k^{app} = E_k^{app}\varphi_k^{app} , \quad E_0^{app} = \left(\varphi_0^{app}, \hat{H}_{app}\varphi_0^{app} \right) .$$

Similarly to the case of the exact Hamiltonian,

$$\hat{F}(\hat{H}_{app}) \simeq E_0^{app} \quad (T \rightarrow 0)$$

and

$$\langle \hat{A} \rangle_{app} \simeq \left(\varphi_0^{app}, \hat{A}\varphi_0^{app} \right) \quad (T \rightarrow 0) .$$

In that way, the Gibbs-Bogolubov inequality at $T = 0$ reduces to the inequality

$$E_0 \leq \left(\varphi_0^{app}, \hat{H}\varphi_0^{app} \right) , \quad (110)$$

which is called the *Rayleigh-Ritz inequality* or the variational principle.

4.8 Bogolubov inequalities

Bogolubov has proved several more useful inequalities [30, 48, 49]. One of them concerns the statistical average of the product of two operators $\langle \hat{A}\hat{B} \rangle$, with a Hamiltonian H . Generally, this average $\langle \hat{A}\hat{B} \rangle$ can be treated as a bilinear form, that is a form linear in both \hat{A} and \hat{B} . The required conditions are $\langle \hat{A}^+\hat{A} \rangle \geq 0$ and $\langle \hat{A}\hat{B} \rangle^* = \langle \hat{B}^+\hat{A}^+ \rangle$. For statistical averages, these conditions are always valid. Note that $\hat{A}^+\hat{A}$ is a semi-positive operator.

Bogolubov inequality for products. The bilinear form $\langle \hat{A}\hat{B} \rangle$, satisfies the inequality

$$\left| \langle \hat{A}\hat{B} \rangle \right|^2 \leq \langle \hat{A}\hat{A}^+ \rangle \langle \hat{B}^+\hat{B} \rangle . \quad (111)$$

Proof. Let us consider the form

$$\langle (x\hat{B}^+ + y\hat{A}) (x^*\hat{B} + y^*\hat{A}^+) \rangle \geq 0 ,$$

which is non-negative for arbitrary complex numbers x and y . This can be rewritten as

$$|x|^2 \langle \hat{B}^+\hat{B} \rangle + xy^* \langle \hat{B}^+\hat{A}^+ \rangle + x^*y \langle \hat{A}\hat{B} \rangle + |y|^2 \langle \hat{A}\hat{A}^+ \rangle \geq 0 .$$

Setting

$$x = - \langle \hat{A}\hat{B} \rangle , \quad y = \langle \hat{B}^+\hat{B} \rangle$$

yields

$$\langle \hat{A}\hat{A}^+ \rangle \langle \hat{B}^+\hat{B} \rangle^2 - \left| \langle \hat{A}\hat{B} \rangle \right|^2 \langle \hat{B}^+\hat{B} \rangle \geq 0 .$$

If $\langle \hat{B}^+ \hat{B} \rangle \neq 0$, then the Bogolubov inequality follows.

And if $\langle \hat{B}^+ \hat{B} \rangle = 0$, then we set $x = z \langle \hat{A} \hat{B} \rangle$ and $y = 1$, with z being a real-valued number. This gives

$$2z \left| \langle \hat{A} \hat{B} \rangle \right|^2 + \langle \hat{A} \hat{A}^+ \rangle \geq 0 .$$

Let us send $z \rightarrow -\infty$. Then the left-hand side tends to $-\infty$, if $\langle \hat{A} \hat{B} \rangle$ is nonzero. This is impossible. Hence $\langle \hat{A} \hat{B} \rangle = 0$ when $\langle \hat{B}^+ \hat{B} \rangle = 0$. The Bogolubov inequality again holds true. \square

Another inequality has been proved by Bogolubov for the average of the commutators of two operators \hat{A} and \hat{B} .

Bogolubov inequality for commutators. For any two operators \hat{A} and \hat{B} the inequality holds:

$$\left| \langle [\hat{A}, \hat{B}] \rangle \right|^2 \leq \frac{\beta}{2} \langle \hat{A}^+ \hat{A} + \hat{A} \hat{A}^+ \rangle \left| \langle [\hat{B}^+, [H, \hat{B}]] \rangle \right| . \quad (112)$$

The proof of this inequality is a bit cumbersome, because of which it is omitted here [30].

4.9 Nernst law

Nernst formulated the law as an empirical fact valid for any equilibrium system. But in quantum statistical mechanics, this law can be proved.

Let $\zeta_0 \geq 1$ be the number of quantum states φ_0 possessing the same energy $E_0 \equiv (\varphi_0, \hat{H} \varphi_0)$ at $T = 0$. This parameter ζ_0 is called the *degeneracy factor*. It is independent from thermodynamic variables, being defined by the quantum properties of the system.

Nernst law. The entropy $S = -\text{Tr } \hat{\rho} \ln \hat{\rho}$ for an equilibrium system at zero-temperature is

$$\lim_{T \rightarrow 0} S = \ln \zeta_0 \geq 0 . \quad (113)$$

Proof. The statistical operator of an equilibrium system has the form $\hat{\rho} = e^{-\beta \hat{H}} / \text{Tr} e^{-\beta \hat{H}}$. When temperature tends to zero, we have

$$\text{Tr } e^{-\beta \hat{H}} \simeq \zeta_0 e^{-\beta E_0} \quad (T \rightarrow 0) .$$

Then limit (113) follows. \square

The quantity $S_0 \equiv \ln \zeta_0$ has the meaning of the Boltzmann entropy. Since the entropy S is an extensive quantity, one has $\lim_{N \rightarrow \infty} S/N < \infty$. Assuming that the degeneracy is finite, $\zeta_0 < \infty$, the quantity S_0 can be neglected in the thermodynamic limit when

$$\lim_{N \rightarrow \infty} \frac{S_0}{N} = \lim_{N \rightarrow \infty} \frac{\ln \zeta_0}{N} = 0 .$$

But S_0 cannot be neglected if $\zeta_0 \propto e^{\alpha N}$, with $\alpha > 0$.

From the Nernst law it follows that the specific heat at zero temperature is zero.

Zero-temperature specific heat. The limit of the specific heat at zero temperature is zero:

$$\lim_{T \rightarrow 0} C_V = 0 . \quad (114)$$

Proof. At temperature tending to zero, one has

$$S_0 = \lim_{T \rightarrow 0} S = \lim_{T \rightarrow 0} \frac{TS}{T} = \lim_{T \rightarrow 0} \frac{\partial(TS)}{\partial T} .$$

Using the expression of the specific heat

$$C_V = \frac{T}{N} \left(\frac{\partial S}{\partial T} \right)_V = \frac{1}{N} \left(\frac{\partial E}{\partial T} \right)_V ,$$

we come to the equality

$$S_0 = S_0 + N \lim_{T \rightarrow 0} C_V .$$

From here, limit (114) follows. \square

It is important to emphasize that, for any degeneracy factor ζ_0 , the limits

$$\lim_{T \rightarrow 0} F = E_0 , \quad \lim_{T \rightarrow 0} \langle \hat{A} \rangle = (\varphi_0, \hat{A} \varphi_0)$$

are valid, since ζ_0 does not depend on T , and

$$\lim_{T \rightarrow 0} T \ln \zeta_0 = 0 .$$

Knowing the behavior of the free energy at low temperature makes it possible to derive an equation for the free energy at arbitrary temperatures. To this end, because of the relations

$$F = E - TS , \quad S = - \frac{\partial F}{\partial T} ,$$

we have the limits

$$\lim_{T \rightarrow 0} F = E_0 , \quad \lim_{T \rightarrow 0} \frac{\partial F}{\partial T} = -S_0 ,$$

serving as boundary conditions for the equation

$$T \frac{\partial F}{\partial T} = F - E .$$

Integrating the latter equation over temperature between zero and T , we obtain the expression

$$F = E_0 - TS_0 - T \int_0^T \frac{E - E_0}{T^2} dT . \quad (115)$$

If S_0 can be neglected in the thermodynamic limit, then the expression for the free energy at large $N \gg 1$ becomes

$$F = E_0 - T \int_0^T \frac{E - E_0}{T^2} dT . \quad (116)$$

The integral over T is well defined because the energy E , as $T \rightarrow 0$, tends to E_0 not slower than as T^2 .

4.10 Parameter variation

It often happens that the Hamiltonian $\hat{H} = \hat{H}(\lambda)$ depends on a parameter λ , and one needs to find the variation of an observable quantity $\langle \hat{A} \rangle = \text{Tr } \hat{\rho} \hat{A}$ that is the average of an operator $\hat{A} = \hat{A}(\lambda)$, which can also depend on this parameter. When the Hamiltonian is a function of a parameter, the statistical operator is the function

$$\hat{\rho} = \frac{e^{-\beta \hat{H}}}{\text{Tr} e^{-\beta \hat{H}}} = \hat{\rho}(\lambda)$$

of this parameter as well.

Before considering the variation of the parameter λ , let us define the *covariance* of two operators \hat{A} and \hat{B} :

$$\text{cov}(\hat{A}, \hat{B}) \equiv \frac{1}{2} \langle [\hat{A}, \hat{B}]_+ \rangle - \langle \hat{A} \rangle \langle \hat{B} \rangle ,$$

where the notation for the anticommutator $[\hat{A}, \hat{B}]_+ \equiv \hat{A}\hat{B} + \hat{B}\hat{A}$ is used. The covariance is symmetric,

$$\text{cov}(\hat{A}, \hat{B}) = \text{cov}(\hat{B}, \hat{A}) ,$$

and additive,

$$\text{cov}(\hat{A} + \hat{B}, \hat{C}) = \text{cov}(\hat{A}, \hat{C}) + \text{cov}(\hat{B}, \hat{C}) .$$

The diagonal covariance is the variance, i.e., the dispersion that for a self-adjoint operator is

$$\text{cov}(\hat{A}, \hat{A}) = \langle \hat{A}^2 \rangle - \langle \hat{A} \rangle^2 = \text{var}(\hat{A}) .$$

For a function of an operator $f = f(\hat{A})$, the derivative with respect to the parameter is defined as the symmetrized derivative

$$\frac{\partial}{\partial \lambda} f(\hat{A}) \equiv \frac{1}{2} \left[\frac{\partial f}{\partial \hat{A}}, \frac{\partial \hat{A}}{\partial \lambda} \right]_+ .$$

Direct differentiation yields

$$\frac{\partial}{\partial \lambda} \langle \hat{A} \rangle = \langle \frac{\partial \hat{A}}{\partial \lambda} \rangle - \beta \text{cov} \left(\hat{A}, \frac{\partial \hat{H}}{\partial \lambda} \right) . \quad (117)$$

Considering the zero-temperature limit, we have the following properties.

Proposition 1. Let \hat{A} be a self-adjoint integral of motion, then

$$\lim_{T \rightarrow 0} \beta \text{cov} \left(\hat{A}, \frac{\partial \hat{H}}{\partial \lambda} \right) = 0 . \quad (118)$$

Proof. A self-adjoint operator is such for which $(\varphi_k, \hat{A}\varphi_p) = (\hat{A}\varphi_k, \varphi_p)$ for any φ_k, φ_p from the Fock space. An operator is the integral of motion when $[\hat{A}, \hat{H}] = 0$. In that case, it enjoys

a common set of eigenfunctions with the Hamiltonian, so that $\hat{A}\varphi_k = A_k\varphi_k$ and $\hat{H}\varphi_k = E_k\varphi_k$. For the normalized functions, such that $(\varphi_k, \varphi_k) = 1$, one has $\partial(\varphi_k, \varphi_k)/\partial\lambda = 0$. Therefore

$$\frac{\partial}{\partial\lambda} (\varphi_k, \hat{A}\varphi_k) = \left(\varphi_k, \frac{\partial\hat{A}}{\partial\lambda} \varphi_k \right) .$$

When $T \rightarrow 0$, then $\lim_{T \rightarrow 0} \langle \hat{A} \rangle = (\varphi_0, \hat{A}\varphi_0)$, with φ_0 being the ground state. In that way,

$$\lim_{T \rightarrow 0} \frac{\partial}{\partial\lambda} \langle \hat{A} \rangle = \lim_{T \rightarrow 0} \left\langle \frac{\partial\hat{A}}{\partial\lambda} \right\rangle .$$

Then, in view of equality (117), limit (118) follows. \square

Proposition 2. For the derivative of the entropy S with respect to a parameter λ , the zero-temperature limit

$$\lim_{T \rightarrow 0} T \frac{\partial S}{\partial\lambda} = 0 \quad (119)$$

is valid.

Proof. Let us take in Eq. (117) $\hat{A} = \hat{H}$. Then for $E \equiv \langle \hat{H} \rangle$, at any $T \geq 0$, we have

$$\frac{\partial E}{\partial\lambda} = \left\langle \frac{\partial\hat{H}}{\partial\lambda} \right\rangle - \beta \operatorname{cov} \left(\hat{H}, \frac{\partial\hat{H}}{\partial\lambda} \right) .$$

For the normalized functions,

$$\frac{\partial}{\partial\lambda} (\varphi_k, \hat{H}\varphi_k) = \left(\varphi_k, \frac{\partial\hat{H}}{\partial\lambda} \varphi_k \right) .$$

According to property (118), one has

$$\lim_{T \rightarrow 0} \beta \operatorname{cov} \left(\hat{H}, \frac{\partial\hat{H}}{\partial\lambda} \right) = 0 .$$

It is straightforward to check that

$$\left\langle \frac{\partial\hat{H}}{\partial\lambda} \right\rangle = \frac{\partial F}{\partial\lambda} , \quad F = -T \ln \operatorname{Tre}^{-\beta\hat{H}} .$$

This leads to the relation

$$\frac{\partial E}{\partial\lambda} = \frac{\partial F}{\partial\lambda} - \beta \operatorname{cov} \left(\hat{H}, \frac{\partial\hat{H}}{\partial\lambda} \right) .$$

Since $F = E - TS$, we get

$$\frac{\partial F}{\partial\lambda} = \frac{\partial E}{\partial\lambda} - T \frac{\partial S}{\partial\lambda} , \quad \frac{\partial E}{\partial\lambda} = \left\langle \frac{\partial\hat{H}}{\partial\lambda} \right\rangle + T \frac{\partial S}{\partial\lambda} .$$

Comparing the above formulas, we obtain

$$\frac{\partial S}{\partial \lambda} = -\beta^2 \text{cov} \left(\hat{H}, \frac{\partial \hat{H}}{\partial \lambda} \right),$$

which, according to Eq. (118), gives limit (119).

Note that the same result can be derived from the expression $S \equiv -\text{Tr} \hat{\rho} \ln \hat{\rho}$, from which

$$\frac{\partial S}{\partial \lambda} = \beta \text{cov} \left(\ln \hat{\rho}, \frac{\partial \hat{H}}{\partial \lambda} \right).$$

Substituting here $\ln \hat{\rho} = \beta(F - \hat{H})$ and taking into account Eq. (118) yields limit (119). \square

Proposition 3. Let the Hamiltonian be the sum $\hat{H} = \hat{K} + \hat{W}$ of the kinetic-energy operator and potential-energy operator, respectively. Then the variation of the internal energy E with respect to mass m , at zero temperature, gives

$$\frac{\partial E}{\partial m} = -\frac{1}{m} \langle \hat{K} \rangle + \langle \frac{\partial \hat{W}}{\partial m} \rangle \quad (T = 0). \quad (120)$$

Proof. Let us set $\lambda = m$. From the form of the kinetic-energy operator, we have

$$\frac{\partial \hat{K}}{\partial m} = -\frac{1}{m} \hat{K}, \quad \frac{\partial \hat{H}}{\partial m} = -\frac{1}{m} \hat{K} + \frac{\partial \hat{W}}{\partial m}.$$

This yields

$$\frac{\partial E}{\partial m} = -\frac{1}{m} \langle \hat{K} \rangle + \langle \frac{\partial \hat{W}}{\partial m} \rangle + T \frac{\partial S}{\partial m},$$

which is equivalent to

$$\frac{\partial F}{\partial m} = -\frac{1}{m} \langle \hat{K} \rangle + \langle \frac{\partial \hat{W}}{\partial m} \rangle.$$

According to Proposition 2, we have

$$\frac{\partial S}{\partial m} = -\beta^2 \text{cov} \left(\hat{H}, \frac{\partial \hat{H}}{\partial m} \right), \quad \lim_{T \rightarrow 0} T \frac{\partial S}{\partial m} = 0.$$

Thus we come to Eq. (120). \square

Equation (120) shows that the Landau-Lifshitz [28] relation $\langle \hat{K} \rangle = -m \partial E / \partial m$, at $T = 0$, is valid only when $\partial \hat{W} / \partial m = 0$. However the latter is not always true. For example, in the case of a dilute system of cold atoms, one has $\partial \hat{W} / \partial m = -\hat{W} / m$. Then

$$\frac{\partial \hat{H}}{\partial m} = -\frac{1}{m} \hat{H}, \quad \frac{\partial S}{\partial m} = \frac{\beta^2}{m} \text{var}(\hat{H}).$$

In view of the expression for the specific heat

$$C_V \equiv \frac{1}{N} \left(\frac{\partial E}{\partial T} \right)_V = \frac{\text{var}(\hat{H})}{NT^2},$$

we get

$$\frac{\partial S}{\partial m} = \frac{N}{m} C_V, \quad \lim_{T \rightarrow 0} T C_V = 0.$$

Therefore, instead of the Landau-Lifshitz relation, we have

$$\frac{\partial E}{\partial m} = - \frac{E}{m} \quad (T = 0). \quad (121)$$

This shows that one has to be cautious when invoking equations presumed to be well known, not forgetting to check the region of applicability of these equations.

4.11 Kubo formula

When the system is subject to an external perturbation, its observable quantities change. Such a change can be described by the Kubo formula [57, 58].

Let the system, at time $t = 0$, be characterized by a Hamiltonian H , and a statistical operator $\hat{\rho}(t)$. Generally, H is a grand Hamiltonian. In the case when there are no external perturbations, the statistical operator satisfies the Liouville equation

$$i \frac{d}{dt} \hat{\rho}(t) = [H, \hat{\rho}(t)].$$

The statistical operator evolution can be represented as $\hat{\rho}(t) = \hat{U}(t) \hat{\rho}(t_0) \hat{U}^\dagger(t)$ by means of the evolution operator satisfying the Schrödinger equation

$$i \frac{d}{dt} \hat{U}(t) = H \hat{U}(t),$$

with the initial condition $\hat{U}(t_0) = 1$. The observable quantity, corresponding to an operator $\hat{A}(t) = \hat{U}^\dagger(t) \hat{A}(t_0) \hat{U}(t)$, when there is no perturbation, is given by the average

$$\langle \hat{A}(t) \rangle \equiv \text{Tr} \hat{\rho}(t) \hat{A}(t_0) = \text{Tr} \hat{\rho}(t_0) \hat{A}(t).$$

But let us assume that at the moment $t \geq t_0$, the Hamiltonian suddenly changes to H_{out} . Hence now the statistical operator becomes $\hat{\rho}_{out}(t)$ that satisfies the Liouville equation

$$i \frac{d}{dt} \hat{\rho}_{out}(t) = [H_{out}, \hat{\rho}_{out}(t)],$$

with the initial condition $\hat{\rho}_{out}(t_0) = \hat{\rho}(t_0)$. Now, the observable quantity, corresponding to the operator \hat{A} , reads as

$$\langle \hat{A}(t) \rangle_{out} \equiv \text{Tr} \hat{\rho}_{out}(t) \hat{A}(t_0).$$

The change of the observable quantity, called *increment*, is defined as

$$\Delta \langle \hat{A}(t) \rangle \equiv \langle \hat{A}(t) \rangle_{out} - \langle \hat{A}(t) \rangle. \quad (122)$$

Let the new Hamiltonian have the form $H_{out} = H + \Delta H$. Then the Liouville equation becomes

$$i \frac{d}{dt} \hat{\rho}_{out}(t) = [H, \hat{\rho}_{out}(t)] + [\Delta H, \hat{\rho}_{out}(t)].$$

Looking for the solution of the Liouville equation in the form

$$\hat{\rho}_{out}(t) = \hat{U}(t)\hat{D}(t)\hat{U}^+(t)$$

and introducing the notation

$$\Delta H(t) \equiv \hat{U}^+(t)\Delta H\hat{U}(t) ,$$

we obtain the equation

$$i \frac{d}{dt} \hat{D}(t) = [\Delta H(t), \hat{D}(t)] ,$$

with the initial condition $\hat{D}(t_0) = \hat{\rho}_{out}(t_0) = \hat{\rho}(t_0)$. The above equation can be rewritten as the integral equation

$$\hat{D}(t) = \hat{\rho}(t_0) - i \int_{t_0}^t [\Delta H(t_1), \hat{D}(t_1)] dt_1 . \quad (123)$$

Using this, the new statistical operator can be represented as

$$\hat{\rho}_{out}(t) = \hat{\rho}(t) - i \hat{U}(t) \int_{t_0}^t [\Delta H(t_1), \hat{D}(t_1)] dt_1 \hat{U}^+(t) . \quad (124)$$

Then increment (122) takes the form

$$\Delta < \hat{A}(t) > = -i \int_{t_0}^t \text{Tr}[\hat{A}(t), \Delta H(t_1)] \hat{D}(t_1) dt_1 . \quad (125)$$

This expression is exact for any ΔH .

By constructing an iterative solution for $\hat{D}(t)$, we start with the zero approximation $\hat{D}_0(t) = \hat{\rho}(t_0)$. Then the first approximation is

$$\hat{D}_1(t) = \hat{D}_0(t) - i \int_{t_0}^t [\Delta H(t_1), \hat{\rho}(t_0)] dt_1 .$$

The second iteration gives

$$\hat{D}_2(t) = \hat{D}_1(t) + (-i)^2 \int_{t_0}^t \int_{t_0}^{t_1} [\Delta H(t_1), [\Delta H(t_2), \hat{\rho}(t_0)]] dt_1 dt_2 .$$

Continuing the iteration procedure, we come to the formal solution

$$\begin{aligned} \hat{D}(t) = & \hat{\rho}(t_0) + \sum_{n=1}^{\infty} (-i)^n \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 \dots \int_{t_0}^{t_{n-1}} dt_n \times \\ & \times [\Delta H(t_1), [\Delta H(t_2), \dots [\Delta H(t_n), \hat{\rho}(t_0)]] \dots] . \end{aligned}$$

Then increment (125) becomes

$$\Delta < \hat{A}(t) > = -i \int_{t_0}^t < [\hat{A}(t), \Delta H(t_1)] > dt_1 + \sum_{n=1}^{\infty} (-i)^{n+1} \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 \dots \int_{t_0}^{t_n} dt_{n+1} \times \quad (126)$$

$$\times < \left[\left[\dots \left[\hat{A}(t), \Delta H(t_1) \right], \Delta H(t_2) \right], \dots, \Delta H(t_{n+1}) \right] > .$$

If the Hamiltonian perturbation $\Delta H \equiv \delta H$ is small, such that $\delta H \rightarrow 0$, then one can limit himself/herself by the first-order approximation for increment (126). Then $\Delta < \hat{A} > \equiv \delta < \hat{A} >$ is termed the *linear response* or the *linear increment* that is

$$\delta < \hat{A}(t) > = -i \int_{t_0}^t < \left[\hat{A}(t), \delta H(t_1) \right] > dt_1 , \quad (127)$$

which is the *Kubo formula*.

A particular important case is when $t_0 \rightarrow -\infty$, which is called the *adiabatic switching*. In that case, the Kubo formula (127) can be represented as

$$\delta < \hat{A}(t) > = -i \int_{-\infty}^{\infty} \Theta(t - t_1) < \left[\hat{A}(t), \delta H(t_1) \right] > dt_1 , \quad (128)$$

where $\Theta(t)$ is a unit step function

$$\Theta(t) = \begin{cases} 1, & t \geq 0 \\ 0, & t < 0 . \end{cases}$$

Let us introduce the *commutator Green function*

$$C \left[\hat{A}(t), \hat{B}(t') \right] \equiv -i \Theta(t - t') < \left[\hat{A}(t), \hat{B}(t') \right] > , \quad (129)$$

which sometimes is named the Bogolubov-Tyablikov Green function. Then the linear response (127) reads as

$$\delta < \hat{A}(t) > = \int_{-\infty}^{\infty} C \left[\hat{A}(t), \delta H(t_1) \right] dt_1 . \quad (130)$$

If one assumes that the system at $t \leq t_0$ has been in equilibrium, with a Hamiltonian H and a statistical operator $\hat{\rho}$, then the commutator Green function (129) depends only on the difference of times $t - t'$. This makes it possible to define the Fourier transform

$$C \left[\hat{A}(t), \hat{B}(t') \right] = \int_{-\infty}^{\infty} \chi \left(\hat{A}, \hat{B}, \omega \right) e^{-i\omega(t-t')} \frac{d\omega}{2\pi} .$$

Respectively, the inverse Fourier transform, called the *admittance*, reads as

$$\chi \left(\hat{A}, \hat{B}, \omega \right) = \int_{-\infty}^{\infty} C \left[\hat{A}(t), \hat{B}(0) \right] e^{i\omega t} dt , \quad (131)$$

where the equality

$$\int_{-\infty}^{\infty} e^{i\omega t} dt = 2\pi \delta(\omega)$$

is used.

Finally, we come to the expression for the linear response

$$\delta < \hat{A}(t) > = \lim_{\omega \rightarrow 0} \chi(\hat{A}, \delta H, \omega) . \quad (132)$$

Here all averages are calculated for an equilibrium system.

4.12 Quasi-equilibrium states

In section 4.2, it has been mentioned that, using the local information functional, one can define a quasi-equilibrium, or locally equilibrium, statistical operator. Here, we shall specify the general properties of the corresponding quasi-equilibrium system [46].

For any system, the energy Hamiltonian and the number-of-particle operator can be written as integrals of the related densities,

$$\hat{H} = \int \hat{H}(x) dx, \quad \hat{N} = \int \hat{n}(x) dx,$$

where $\hat{n}(x) \equiv \psi^\dagger(x)\psi(x)$ is the density-of-particle operator. One can introduce the local energy density $E(x) = \langle \hat{H}(x) \rangle$ and the local particle density $\rho(x) = \langle \hat{n}(x) \rangle$. Their integrals define the internal energy and the number of particles,

$$E = \int E(x) dx, \quad N = \int \rho(x) dx.$$

The local information functional can be written in the form

$$\begin{aligned} I_{loc}[\hat{\rho}] = & \text{Tr} \hat{\rho} \ln \hat{\rho} + \lambda_0 (\text{Tr} \hat{\rho} - 1) + \int \lambda_1(x) [\text{Tr} \hat{\rho} \hat{H}(x) - E(x)] dx + \\ & + \int \lambda_2(x) [\text{Tr} \hat{\rho} \hat{n}(x) - \rho(x)] dx, \end{aligned}$$

with the Lagrange multipliers guaranteeing the normalization condition and the definitions of the local energy density and local particle density. These multipliers can be rewritten as $\lambda_0 \equiv \ln Z - 1$, $\lambda_1(x) \equiv \beta(x)$, and $\lambda_2(x) \equiv -\beta(x)\mu(x)$, where $\beta(x)$ is the local inverse temperature, such that $\beta(x)T(x) = 1$, with $T(x)$ being the local temperature, and $\mu(x)$ is the local chemical potential.

From the principle of minimal information, we find the local statistical operator

$$\hat{\rho} = \frac{1}{Z} \exp \left\{ - \int \beta(x) [\hat{H}(x) - \mu(x)\hat{n}(x)] dx \right\}, \quad (133)$$

with the local partition function

$$Z = \text{Tr} \exp \left\{ - \int \beta(x) [\hat{H}(x) - \mu(x)\hat{n}(x)] dx \right\}.$$

The *quasi-equilibrium potential* $Q \equiv -\ln Z$ can be represented through the *local grand potential* $\Omega(x)$, so that

$$Q = \int \beta(x)\Omega(x) dx, \quad Z \equiv \exp \left\{ - \int \beta(x)\Omega(x) dx \right\}. \quad (134)$$

From $Q = -\ln Z$, we have

$$\frac{\delta Q}{\delta \beta(x)} = E(x) - \mu(x)\rho(x).$$

While from the first relation with $\Omega(x)$ in Eq. (134), we get

$$\frac{\delta Q}{\delta \beta(x)} = \Omega(x) + \beta(x) \frac{\delta \Omega(x)}{\delta \beta(x)} .$$

Comparing the above formulas, we find

$$\frac{\delta \Omega(x)}{\delta \beta(x)} = T(x) [E(x) - \mu(x)\rho(x) - \Omega(x)] .$$

Because of the equality

$$\frac{\delta \Omega(x)}{\delta \beta(x)} = -T^2(x) \frac{\delta \Omega(x)}{\delta T(x)} ,$$

we obtain

$$\frac{\delta \Omega(x)}{\delta T(x)} = \beta(x) [\Omega(x) + \mu(x)\rho(x) - E(x)] .$$

The *local entropy* is introduced by the definition

$$S(x) \equiv - \frac{\delta \Omega(x)}{\delta T(x)} ,$$

which results in the expression

$$S(x) = \beta(x) [E(x) - \mu(x)\rho(x) - \Omega(x)] . \quad (135)$$

From here, we have the relations

$$\beta(x) = \frac{\delta S(x)}{\delta E(x)} , \quad \mu(x) = -T(x) \frac{\delta S(x)}{\delta \rho(x)} .$$

The Gibbs entropy (information entropy) can be represented as an integral of the local entropy:

$$S = -\text{Tr} \hat{\rho} \ln \hat{\rho} = \int S(x) dx .$$

Here the local entropy is as defined in Eq. (135), since

$$S = \int \beta(x) [E(x) - \mu(x)\rho(x)] dx - Q .$$

Taking into account the equalities

$$\frac{\delta Q}{\delta \mu(x)} = -\beta(x)\rho(x) , \quad \frac{\delta Q}{\delta \mu(x)} = \beta(x) \frac{\delta \Omega(x)}{\delta \mu(x)} , \quad \rho(x) = - \frac{\delta \Omega(x)}{\delta \mu(x)} ,$$

we come to the *local Gibbs-Duhem relation*

$$\delta \Omega(x) = -S(x)\delta T(x) - \rho(x)\delta \mu(x) . \quad (136)$$

Introducing the local pressure $P(x) \equiv -\Omega(x)$ yields

$$\delta P(x) = S(x)\delta T(x) + \rho(x)\delta \mu(x) . \quad (137)$$

The local quantities naturally appear in nonuniform systems, such as the systems of trapped atoms.

4.13 Gauge symmetry

One of the most important properties of systems is gauge symmetry. When the system is gauge symmetric, there can be neither Bose-Einstein condensation nor superconductivity.

The typical energy Hamiltonian is

$$\begin{aligned} \hat{H} = & \int \psi^\dagger(x) \left[-\frac{\nabla^2}{2m} + U(x) \right] \psi(x) dx + \\ & + \frac{1}{2} \int \psi^\dagger(x) \psi^\dagger(x') \Phi(x, x') \psi(x') \psi(x) dx dx' , \end{aligned} \quad (138)$$

where $U(x)$ is an external potential and $\Phi(x, x')$ is a pair interaction potential.

The number-of-particle operator is an integral of motion,

$$[\hat{N}, \hat{H}] = 0 , \quad \hat{N} = \int \psi^\dagger(x) \psi(x) dx .$$

This is termed the number-of-particle conservation.

The *global gauge transformation* is accomplished by the action of the operator

$$\hat{U}_\varphi \equiv e^{i\varphi \hat{N}} , \quad (139)$$

where φ is the global phase, being any real number.

Since $[\hat{N}, \hat{H}] = 0$, Hamiltonian (138) is gauge symmetric:

$$\hat{U}_\varphi^\dagger \hat{H} \hat{U}_\varphi = \hat{H} .$$

In other words, \hat{H} is invariant under the field-operator transformation

$$\psi(x) \rightarrow \hat{U}_\varphi^\dagger \psi(x) \hat{U}_\varphi .$$

The gauge transformation \hat{U}_φ enjoys the following properties:

$$\hat{U}_0 = \hat{1} , \quad \hat{U}_\varphi^{-1} = e^{-i\varphi \hat{N}} , \quad \hat{U}_\varphi^\dagger \hat{U}_\varphi = \hat{1} .$$

It also possesses the property of additivity: $\hat{U}_\varphi \hat{U}_{\varphi'} = \hat{U}_{\varphi+\varphi'}$. Therefore the family $\{\hat{U}_\varphi\}$ forms a group, called the *gauge group*.

Among various averages of operator products, one distinguishes two principally different types of the averages. One type, called *normal averages* consists of the products containing equal numbers of creation and destruction field operators, e.g., as in the average $\langle \psi^\dagger(x) \psi(x') \rangle$. The other type of the averages, termed *anomalous averages* consists of the products with unequal numbers of the creation and destruction operators, for instance, as in the averages $\langle \psi(x) \rangle$ or $\langle \psi(x) \psi(x') \rangle$.

Statement 1. If the number-of-particle operator is an integral of motion, then all anomalous averages are identically zero.

Proof. The particle-number operator has the following properties of commutation with the field operators:

$$[\psi(x), \hat{N}] = \psi(x) , \quad [\psi^\dagger(x), \hat{N}] = -\psi^\dagger(x) .$$

From here, one gets

$$\psi(x)\hat{N} = (1 + \hat{N})\psi(x) , \quad \psi(x)\hat{N}^n = (1 + \hat{N})^n\psi(x) .$$

The gauge-transformation operator (139) can be represented as

$$\hat{U}_\varphi = \sum_{n=-\infty}^{\infty} \frac{(i\varphi)^n}{n!} \hat{N}^n .$$

Using this, we come to the relations

$$\psi(x)\hat{U}_\varphi = e^{i\varphi}\hat{U}_\varphi\psi(x) , \quad \hat{U}_\varphi^+\psi(x)\hat{U}_\varphi = e^{i\varphi}\psi(x) , \quad \hat{U}_\varphi^+\psi^\dagger(x)\hat{U}_\varphi = e^{-i\varphi}\psi^\dagger(x) .$$

In equilibrium, $[\hat{\rho}, \hat{H}] = 0$. Since \hat{N} is an integral of motion, it follows that

$$[\hat{N}, \hat{\rho}] = 0 , \quad [\hat{U}_\varphi, \hat{\rho}] = 0 .$$

Because of the above equalities, we obtain

$$\begin{aligned} \langle \psi(x)\psi(x') \rangle &= \text{Tr} \hat{\rho} \psi(x)\psi(x') = \text{Tr} \hat{U}_\varphi \hat{\rho} \hat{U}_\varphi^+ \psi(x)\psi(x') = \\ &= \text{Tr} \hat{\rho} \hat{U}_\varphi^+ \psi(x)\psi(x') \hat{U}_\varphi = e^{2i\varphi} \text{Tr} \hat{\rho} \psi(x)\psi(x') . \end{aligned}$$

Therefore the equality

$$\langle \psi(x)\psi(x') \rangle = (1 - e^{2i\varphi}) = 0$$

is valid for any φ . This implies that

$$\langle \psi(x)\psi(x') \rangle = 0 , \tag{140}$$

while $\langle \psi^\dagger(x)\psi(x') \rangle \neq 0$. In the same way, it is straightforward to show that

$$\langle \psi(x) \rangle = (1 - e^{i\varphi}) = 0$$

for any φ , hence $\langle \psi(x) \rangle = 0$. Generally, we find the equality

$$\langle \prod_{i=1}^n \psi^\dagger(x_i) \prod_{j=1}^m \psi(x_j) \rangle = [1 - e^{i\varphi(m-n)}] = 0 ,$$

telling us that it is zero for $m \neq n$. Because this is zero for an arbitrary φ , we come to the conclusion that

$$\langle \prod_{i=1}^n \psi^\dagger(x_i) \prod_{j=1}^m \psi(x_j) \rangle = 0 \quad (m \neq n) . \tag{141}$$

Thus, all anomalous averages are identically zero. \square

Contrary to this, the normal averages are not zero,

$$\langle \prod_{i=1}^n \psi^\dagger(x_i) \psi(x'_i) \rangle \neq 0 .$$

Since the appearance of the anomalous averages necessarily accompanies Bose-Einstein condensation or superconducting transition [1,12], this means that these transitions cannot occur in a gauge-symmetric system, but the symmetry breaking must happen.

4.14 Translation invariance

When the operator of momentum is an integral of motion, there exist constraints on the averages, from which it follows that such a system has to be uniform. In this section, we understand the variable x as a spatial variable, $x \rightarrow \mathbf{r}$.

Let us assume that there is no external potential, $U(\mathbf{r}) = 0$. Then the operator of momentum is an integral of motion:

$$[\hat{\mathbf{P}}, \hat{H}] = 0, \quad \hat{\mathbf{P}} = \int \psi^\dagger(\mathbf{r})(-i\nabla)\psi(\mathbf{r}) d\mathbf{r}.$$

This is called the *momentum conservation*. The family of the *translation transformations*

$$\hat{T}(\mathbf{r}) = \exp(i\hat{\mathbf{P}} \cdot \mathbf{r}) \quad (142)$$

forms the translation group $\{\hat{T}(\mathbf{r})\}$. Recall that, according to Sec. 3.8, one has $\psi(\mathbf{r}) = \hat{T}^+(\mathbf{r})\psi(0)\hat{T}(\mathbf{r})$. When there is no external potential, the momentum operator commutes with Hamiltonian (138), so that

$$[\hat{T}(\mathbf{r}), \hat{H}] = 0, \quad [\hat{T}(\mathbf{r}), \hat{\rho}] = 0,$$

for an equilibrium $\hat{\rho}$.

Statement 2. When the operator of momentum is an integral of motion, then

$$\langle \prod_{i=1}^n \psi^\dagger(\mathbf{r}_i + \mathbf{r}) \prod_{j=1}^m \psi(\mathbf{r}_j + \mathbf{r}) \rangle = \langle \prod_{i=1}^n \psi^\dagger(\mathbf{r}_i) \prod_{j=1}^m \psi(\mathbf{r}_j) \rangle, \quad (143)$$

for any integers m, n .

Proof. Starting with the case of two field operators, we have

$$\begin{aligned} \langle \psi^\dagger(\mathbf{r}_1 + \mathbf{r})\psi(\mathbf{r}_2 + \mathbf{r}) \rangle &= \text{Tr} \hat{\rho} \hat{T}^+(\mathbf{r})\psi^\dagger(\mathbf{r}_1)\psi(\mathbf{r}_2)\hat{T}(\mathbf{r}) = \\ &= \text{Tr} \hat{T}^+(\mathbf{r})\hat{\rho}\psi^\dagger(\mathbf{r}_1)\psi(\mathbf{r}_2)\hat{T}(\mathbf{r}) = \text{Tr} \hat{\rho}\psi^\dagger(\mathbf{r}_1)\psi(\mathbf{r}_2). \end{aligned}$$

That is,

$$\langle \psi^\dagger(\mathbf{r}_1 + \mathbf{r})\psi(\mathbf{r}_2 + \mathbf{r}) \rangle = \langle \psi^\dagger(\mathbf{r}_1)\psi(\mathbf{r}_2) \rangle.$$

Similarly, for the product of any number of field operators, we obtain Eq. (143). \square

In particular, we have

$$\langle \psi^\dagger(\mathbf{r}_1)\psi(\mathbf{r}_2) \rangle = \langle \psi^\dagger(\mathbf{r}_1 - \mathbf{r}_2)\psi(0) \rangle,$$

which leads to the uniform particle density

$$\langle \psi^\dagger(\mathbf{r})\psi(\mathbf{r}) \rangle = \langle \psi^\dagger(0)\psi(0) \rangle.$$

For a uniform system of N atoms in volume V , it is convenient to use the Fourier basis $\{\varphi_k(\mathbf{r})\}$ of plane waves

$$\varphi_k(\mathbf{r}) = \frac{1}{\sqrt{V}} e^{i\mathbf{k} \cdot \mathbf{r}}.$$

The basis is orthonormalized, $(\varphi_k, \varphi_p) = \delta_{kp}$, and complete,

$$\sum_k \varphi_k(\mathbf{r}) \varphi_k^*(\mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}') .$$

Imposing the periodic boundary conditions $\varphi_k(\mathbf{r} + \mathbf{L}) = \varphi_k(\mathbf{r})$, where $\mathbf{L} = \{L^\alpha\}$ and $V = \prod_\alpha L^\alpha$, one has $\mathbf{k} \cdot \mathbf{L} = 2\pi$, hence

$$\mathbf{k} = \{k^\alpha\} , \quad k^\alpha = \frac{2\pi}{L^\alpha} n^\alpha \quad (n^\alpha = 0, \pm 1, \pm 2, \dots) .$$

In the thermodynamic limit, when $N \rightarrow \infty, V \rightarrow \infty, L^\alpha \rightarrow \infty$, the difference between the neighbor wave numbers $\Delta k^\alpha = 2\pi/L^\alpha$, for which $\Delta n^\alpha = 1$, tends to zero, $\Delta k^\alpha \rightarrow 0$. In that case, the summation over the wave numbers k^α , which is equivalent to the summation over n^α , so that $\sum_k \equiv \sum_{\{n^\alpha\}}$, can be replaced by the integration

$$\sum_{\{n^\alpha\}} \rightarrow \int d\mathbf{n} \quad (\mathbf{n} = \{n^\alpha\}) .$$

Since

$$dn^\alpha = L^\alpha \frac{dk^\alpha}{2\pi} , \quad d\mathbf{n} = V \frac{d\mathbf{k}}{(2\pi)^3} ,$$

the replacement takes the form

$$\sum_k \rightarrow V \int \frac{d\mathbf{k}}{(2\pi)^3} .$$

Passing to the momentum representation, one expands the field operators over the Fourier basis,

$$\psi(\mathbf{r}) = \sum_k a_k \varphi_k(\mathbf{r}) ,$$

where

$$a_k = (\varphi_k, \psi) = \int \varphi_k^*(\mathbf{r}) \psi(\mathbf{r}) d\mathbf{r} .$$

From the commutation relations $[\psi(\mathbf{r}), \psi(\mathbf{r}')]_{\mp} = 0$ and $[\psi(\mathbf{r}), \psi^\dagger(\mathbf{r}')]_{\mp} = \delta(\mathbf{r} - \mathbf{r}')$, one gets the commutation relations

$$[a_k, a_p]_{\mp} = 0 , \quad [a_k, a_p^\dagger]_{\mp} = \delta_{kp} .$$

Because of the momentum conservation, the system is translation invariant, such that

$$\frac{\partial}{\partial \mathbf{r}} < \psi^\dagger(\mathbf{r}_1 + \mathbf{r}) \psi(\mathbf{r}_2 + \mathbf{r}) > = 0 .$$

Passing to the momentum representation in the average

$$< \psi^\dagger(\mathbf{r}_1 + \mathbf{r}) \psi(\mathbf{r}_2 + \mathbf{r}) > = \sum_{kp} < a_k^\dagger a_p > \varphi_k^*(\mathbf{r}_1 + \mathbf{r}) \varphi_p(\mathbf{r}_2 + \mathbf{r})$$

yields

$$\frac{\partial}{\partial \mathbf{r}} < \psi^\dagger(\mathbf{r}_1 + \mathbf{r}) \psi(\mathbf{r}_2 + \mathbf{r}) > = -i \sum_{kp} < a_k^\dagger a_p > (\mathbf{k} - \mathbf{p}) \varphi_k^*(\mathbf{r}_1 + \mathbf{r}) \varphi_p(\mathbf{r}_2 + \mathbf{r}) = 0 .$$

From here, it is seen that for a uniform system, one has the properties

$$\langle a_k^\dagger a_p \rangle = \delta_{kp} \langle a_k^\dagger a_k \rangle, \quad \langle a_k a_p \rangle = \delta_{-kp} \langle a_{-k} a_k \rangle. \quad (144)$$

In general, the average

$$\langle \prod_{i=1}^n a_{k_i}^\dagger \prod_{j=1}^m a_{p_j} \rangle \neq 0 \quad (145)$$

is not zero only when $\mathbf{k}_1 + \dots + \mathbf{k}_n = \mathbf{p}_1 + \dots + \mathbf{p}_m$.

In this way, if the operator of momentum is an integral of motion, the system is uniform and cannot characterize nonuniform systems, such as crystalline matter. To be able to describe the latter, translation symmetry breaking is necessary.

4.15 Symmetry breaking

As is stressed in the previous two sections, in order to describe some thermodynamic phases, it is necessary to consider systems with broken symmetries. For instance, if the Hamiltonian of a magnetic system is invariant under spin inversion or rotation, that is, exhibits the related symmetry with respect to the magnetic moment \mathbf{M} , then the latter is identically zero [49].

Or, when the operator of momentum is an integral of motion, $[\hat{\mathbf{P}}, \hat{H}] = 0$, the system enjoys translation symmetry, as is shown in Sec. 4.14, because of which the average $\langle \psi^\dagger(\mathbf{r})\psi(\mathbf{r}) \rangle = \langle \psi^\dagger(0)\psi(0) \rangle$ does not depend on the spatial variable. But the average

$$\rho(\mathbf{r}) = \langle \psi^\dagger(\mathbf{r})\psi(\mathbf{r}) \rangle \quad (146)$$

represents the spatial distribution of particles, that is, the particle density. If $\rho(\mathbf{r}) = \rho(0) = \text{const}$, then there can be no crystalline matter.

Similarly, as is discussed in Sec. 4.13, when the system is gauge-symmetric, all anomalous averages are identically zero, hence there can be neither Bose-Einstein condensation nor superconductivity.

The above examples demonstrate that, for characterizing some thermodynamic phases, one has to look for states with a broken symmetry. The general idea of how the symmetry, prescribed by the Hamiltonian, can be broken is straightforward. One has to impose some additional conditions, boundary conditions, initial conditions, or like that, for selecting from the total Fock space the microscopic states possessing the required symmetry properties. This is analogous to the selection of required solutions of nonlinear equations possessing multiple solutions.

The direct way of selecting the required states is by the *method of restricted trace*. One defines a subspace of the Fock space $\mathcal{F}_\nu \in \mathcal{F}$, whose states enjoy the necessary symmetry properties. And, calculating the observable quantities, one takes the trace only over the states of the selected subspace,

$$\langle \hat{A} \rangle_\nu \equiv \text{Tr}_{\mathcal{F}_\nu} \hat{\rho} \hat{A}. \quad (147)$$

A general method of symmetry breaking has been developed by Bogolubov [49] and is called the *method of Bogolubov quasiaverages*. The idea of the method is as follows. Let the system Hamiltonian \hat{H} possess some symmetry related to an integral of motion \hat{C} , so that $[\hat{C}, \hat{H}] = 0$. Suppose we wish to break the corresponding symmetry. For this purpose, we introduce the Hamiltonian

$$\hat{H}_\nu(\varepsilon) \equiv \hat{H} + \varepsilon \hat{\Gamma}_\nu, \quad (148)$$

adding to \hat{H} a term $\hat{\Gamma}_\nu$, called *source*, that breaks the symmetry because it does not commute with the integral of motion \hat{C} ,

$$[\hat{C}, \hat{\Gamma}_\nu] \neq 0, \quad [\hat{C}, \hat{H}_\nu(\varepsilon)] \neq 0.$$

The average of the source $\langle \Gamma_\nu \rangle$ is assumed to be an extensive quantity proportional to N for large $N \gg 1$. The index ν enumerates the admissible phases.

The *quasiaverage* of an operator $\hat{A}(\mathbf{r})$ is

$$\langle \hat{A}(\mathbf{r}) \rangle_\nu \equiv \lim_{\varepsilon \rightarrow 0} \lim_{N \rightarrow \infty} \langle \hat{A}(\mathbf{r}) \rangle_{H_\nu(\varepsilon)}, \quad (149)$$

where

$$\langle \hat{A} \rangle_H \equiv \frac{\text{Tr} \hat{A} e^{-\beta \hat{H}}}{\text{Tr} e^{-\beta \hat{H}}}.$$

Here $\lim_{N \rightarrow \infty}$ implies the thermodynamic limit, when $N \rightarrow \infty$, $V \rightarrow \infty$ and $N/V \rightarrow \text{const}$. It is principal that the limits $\varepsilon \rightarrow 0$ and $N \rightarrow \infty$ do not commute.

Instead of the double limit, as in the Bogolubov method, it is possible to employ a single limit, as in the *method of thermodynamic quasiaverages* [59], by introducing the Hamiltonian

$$\hat{H}_\nu \equiv H + \frac{1}{N^\alpha} \hat{\Gamma}_\nu \quad (0 < \alpha < 1), \quad (150)$$

where the restriction $0 < \alpha < 1$ guarantees that the factor $1/N^\alpha$, in the thermodynamic limit, goes to zero slower than $1/N$, which is necessary for the efficient selection of the states with broken symmetry. The average of the source $\langle \Gamma_\nu \rangle$ is again assumed to be an extensive quantity proportional to N for large $N \gg 1$. Then the *thermodynamic quasiaverage* is defined by the sole thermodynamic limit

$$\langle \hat{A}(\mathbf{r}) \rangle_\nu \equiv \lim_{N \rightarrow \infty} \langle \hat{A}(\mathbf{r}) \rangle_{H_\nu}. \quad (151)$$

For example, to break the spin rotational symmetry, it is sufficient to add to the Hamiltonian the source representing the action of an external magnetic field. To break the translation symmetry, which is necessary for describing crystalline matter, it is sufficient to add the source

$$\hat{\Gamma}_\nu = \int \psi^\dagger(\mathbf{r}) U(\mathbf{r}) \psi(\mathbf{r}) d\mathbf{r}, \quad (152)$$

in which the external potential is periodic over the lattice vectors $\{\mathbf{a}\}$, such that $U(\mathbf{r} + \mathbf{a}) = U(\mathbf{r})$ and $[\hat{\mathbf{P}}, \hat{\Gamma}_\nu] \neq 0$. Then the particle density, defined as the Bogolubov quasiaverage

$$\rho_\nu(\mathbf{r}) \equiv \lim_{\varepsilon \rightarrow 0} \lim_{N \rightarrow \infty} \langle \psi^\dagger(\mathbf{r}) \psi(\mathbf{r}) \rangle_{H_\nu(\varepsilon)}, \quad (153)$$

where the index ν corresponds to the chosen crystalline lattice vectors $\{\mathbf{a}\}$, is periodic over the lattice vectors,

$$\rho_\nu(\mathbf{r} + \mathbf{a}) = \rho_\nu(\mathbf{r}). \quad (154)$$

Or one can use the thermodynamic quasiaverages (151) defining the single limit

$$\rho_\nu(\mathbf{r}) \equiv \lim_{N \rightarrow \infty} \langle \psi^\dagger(\mathbf{r}) \psi(\mathbf{r}) \rangle_{H_\nu}. \quad (155)$$

The method of breaking the translation symmetry by adding a source with an external potential that is periodic over the considered crystalline lattice was, first, advanced by Kirkwood [47]. Bogolubov [49] extended this method to the case of an arbitrary symmetry and gave it a precise mathematical meaning. The replacement of the double limit in the Bogolubov method by a single thermodynamic limit was suggested in Ref. [59]. Some other methods of symmetry breaking are reviewed in Ref. [33].

Generally, there can exist several equilibrium states of a system, under the same thermodynamic variables. However, not all those states are stable. The system always chooses a state that is the most stable. When it happens that a state with a broken symmetry is more stable than the symmetric state, the system passes to that more stable state with broken symmetry. This is termed *spontaneous symmetry breaking*.

5 Quadratic Hamiltonians

Averages can be calculated exactly, when the system Hamiltonian is quadratic with respect to field operators. However, such a trivial situation happens rather rarely, corresponding to ideal gases. Interacting systems, as a rule, do not allow for exact calculations. But approximate calculations are possible by reducing the exact Hamiltonian to an approximate quadratic form.

5.1 Occupation numbers

In many cases, it is convenient to invoke the representation of occupation numbers. This representation is introduced as follows. Let us, first, pass to the quantum-number representation by expanding the field operator over an orthonormal basis,

$$\psi(x) = \sum_k a_k \varphi_k(x) , \quad (156)$$

where the operators a_k and a_k^\dagger generate the Fock space $\mathcal{F}(a_k)$, with k being a set of quantum numbers (multi-index). Depending on the particle statistics, the commutation relations are

$$[a_k, a_p]_{\mp} = 0 , \quad [a_k, a_p^\dagger]_{\mp} = \delta_{kp} ,$$

the upper sign corresponding to Bose statistics, while the lower, to Fermi statistics.

The *occupation-number operator* is

$$\hat{n}_k \equiv a_k^\dagger a_k . \quad (157)$$

This is a self-adjoint operator, $\hat{n}_k^\dagger = \hat{n}_k$, and it is semipositive, since, for any $f \in \mathcal{F}(a_k)$, one has

$$(f, \hat{n}_k f) = (f, a_k^\dagger a_k f) = (a_k f, a_k f) = \|a_k f\|^2 \geq 0 .$$

The number-of-particle operator can be expressed as the sum

$$\hat{N} = \int \psi^\dagger(x) \psi(x) dx = \sum_k a_k^\dagger a_k = \sum_k \hat{n}_k$$

of the occupation-number operators.

For any type of statistics, the commutation relations are valid:

$$[\hat{n}_k, \hat{n}_p] = 0, \quad [a_k, \hat{n}_p] = \delta_{kp} a_k, \quad [a_k^\dagger, \hat{n}_p] = -\delta_{kp} a_k^\dagger,$$

where all square brackets represent commutators.

The eigenvalues ν_k of \hat{n}_k , called the *occupation numbers*, are given by the eigenproblem

$$\hat{n}_k |k\rangle = \nu_k |k\rangle. \quad (158)$$

They are non-negative, $\nu_k \geq 0$, since \hat{n}_k is semipositive.

For *Fermi statistics*, the eigenvalues are easily found. As far as $a_k a_k = -a_k a_k$, one has $a_k^2 = 0$. Then

$$\hat{n}_k^2 = a_k^\dagger a_k a_k^\dagger a_k = a_k^\dagger (1 - a_k^\dagger a_k) a_k = a_k^\dagger a_k = \hat{n}_k.$$

This gives $\nu_k^2 = \nu_k$, since $\hat{n}_k^2 |k\rangle = \nu_k^2 |k\rangle$. From here it follows that $\nu_k = 0, 1$, for Fermi statistics.

But in the case of *Bose statistics*, the situation is a bit more involved. First, we notice that if $|k\rangle$ is an eigenvector of \hat{n}_k , then $(a_k)^m |k\rangle$ and $(a_k^\dagger)^m |k\rangle$ are also its eigenvectors. To prove this, let us consider the equalities

$$\hat{n}_k a_k |k\rangle = (a_k \hat{n}_k + [\hat{n}_k, a_k]) |k\rangle = (\nu_k - 1) a_k |k\rangle,$$

$$\hat{n}_k a_k^\dagger |k\rangle = (a_k^\dagger \hat{n}_k + [\hat{n}_k, a_k^\dagger]) |k\rangle = (\nu_k + 1) a_k^\dagger |k\rangle.$$

Continuing further, we find

$$\hat{n}_k (a_k)^m |k\rangle = (\nu_k - m) (a_k)^m |k\rangle, \quad \hat{n}_k (a_k^\dagger)^m |k\rangle = (\nu_k + m) (a_k^\dagger)^m |k\rangle.$$

The eigenvalues of \hat{n}_k cannot be negative, which means that, after some m , the null vector $|0\rangle$ must exist, such that

$$(a_k)^m |k\rangle = |0\rangle, \quad a_k |0\rangle = 0.$$

For the null vector, the eigenvalue of \hat{n}_k is zero. Hence, the eigenvalues ν_k of \hat{n}_k are non-negative integers $\nu_k = 0, 1, 2, \dots$ ranging between 0 and m . The latter can be set to be arbitrarily large.

The occupation-number representation is convenient for the use in the case of quadratic Hamiltonians, such as in the case of ideal gases, or when the Hamiltonian can be approximately reduced to the quadratic form. Then any function $f(\hat{n}_k)$ of the occupation-number operator, acting on the eigenvectors of the latter, yields $f(\hat{n}_k) |k\rangle = f(n_k) |k\rangle$.

5.2 Noninteracting particles

The simplest case of a Hamiltonian, quadratic in the field operators, is that characterizing the ideal gases of noninteracting particles. Let us consider the Gibbs grand ensemble, with the grand Hamiltonian $H \equiv \hat{H} - \mu \hat{N}$. For noninteracting particles, one has

$$H = \int \psi^\dagger(x) \left[-\frac{\nabla^2}{2m} + U(x) - \mu \right] \psi(x) dx, \quad (159)$$

where $U(x)$ is an external potential. The field operators can be expanded, as in Eq. (156), over the orthonormal and complete basis $\{\varphi_k(x)\}$ of the solutions to the Schrödinger equation

$$\left[-\frac{\nabla^2}{2m} + U(x) - \mu \right] \varphi_k(x) = \varepsilon_k \varphi_k(x) .$$

As a result, Hamiltonian (159) reduces to the simple form

$$H = \sum_k (\varepsilon_k - \mu) a_k^\dagger a_k = \sum_k \omega_k a_k^\dagger a_k , \quad (160)$$

with the single-particle spectrum $\omega_k \equiv \varepsilon_k - \mu$.

This Hamiltonian is evidently invariant under the local gauge transformations

$$a_k \rightarrow a_k e^{i\alpha_k} ,$$

where α_k is a real-valued number. Therefore all anomalous averages are identically zero, such as $\langle a_k a_p \rangle = 0$, while for the normal averages, we have $\langle a_k^\dagger a_p \rangle = \delta_{kp} \langle a_k^\dagger a_k \rangle$. The average of the occupation-number operator $\hat{n}_k \equiv a_k^\dagger a_k$ gives the *occupation-number distribution*

$$n_k \equiv \langle \hat{n}_k \rangle . \quad (161)$$

To calculate the distribution n_k , we need to find the partition function $Z = \text{Tr} e^{-\beta H}$. To this end, we notice that

$$\text{Tr} \exp \left(-\beta \sum_k \omega_k a_k^\dagger a_k \right) = \text{Tr} \prod_k \exp(-\beta \omega_k \hat{n}_k) = \prod_k \sum_{\nu_k} \exp(-\beta \omega_k \nu_k) ,$$

with $\nu_k = 0, 1, 2, \dots$ for Bose statistics and $\nu_k = 0, 1$ for Fermi statistics. Hence, for Bose statistics,

$$\sum_{\nu_k} e^{-\beta \omega_k \nu_k} = 1 + e^{-\beta \omega_k} + e^{-2\beta \omega_k} + \dots = \frac{1}{1 - e^{-\beta \omega_k}} .$$

Consequently,

$$Z = \prod_k (1 - e^{-\beta \omega_k})^{-1} \quad (\text{Bose}) .$$

For Fermi statistics,

$$\sum_{\nu_k} e^{-\beta \omega_k \nu_k} = 1 + e^{-\beta \omega_k} .$$

Therefore,

$$Z = \prod_k (1 + e^{-\beta \omega_k}) \quad (\text{Fermi}) .$$

Combining both these cases, we get

$$Z = \prod_k (1 \mp e^{-\beta \omega_k})^{\mp 1} . \quad (162)$$

The grand potential is

$$\Omega = -T \ln Z = -T \ln \text{Tr} e^{-\beta H} = \pm T \sum_k \ln (1 \mp e^{-\beta \omega_k}) .$$

From here, we find the occupation-number distribution

$$n_k = \langle \hat{n}_k \rangle = \frac{\delta \Omega}{\delta \omega_k} = \frac{1}{e^{\beta \omega_k} \mp 1} . \quad (163)$$

Using the identity $e^{\beta \omega_k} = (1 \pm n_k)/n_k$, we get the grand potential

$$\Omega = \mp T \sum_k \ln(1 \pm n_k) . \quad (164)$$

Other thermodynamic quantities can also be calculated exactly using the derivatives

$$\frac{\partial n_k}{\partial T} = \beta^2 \omega_k n_k (1 \pm n_k) , \quad \frac{\partial n_k}{\partial \mu} = \beta n_k (1 \pm n_k) .$$

Thus the internal energy is

$$E = \langle H \rangle + \mu = \sum_k \varepsilon_k n_k .$$

For the entropy, we have

$$S = - \left(\frac{\partial \Omega}{\partial T} \right)_{V\mu} = \sum_k [\beta \omega_k n_k \pm \ln(1 \pm n_k)] .$$

The number of particles is

$$N = - \frac{\partial \Omega}{\partial \mu} = \sum_k n_k .$$

Since $\Omega = -PV$, the pressure becomes

$$P = \pm \frac{T}{V} \sum_k \ln(1 \pm n_k) .$$

The free energy is $F = E - TS = \Omega + \mu N$.

It is worth recalling that \hat{n}_k is a semipositive operator, hence the distribution $n_k = \langle \hat{n}_k \rangle$ has also to be semipositive. This implies that $e^{\beta \omega_k} \geq \pm 1$. For Fermi statistics, $n_k \geq 0$ in any case. But for Bose statistics, it should be $e^{\beta \omega_k} \geq 1$, which means that $\omega_k \geq 0$, so that $\mu \leq \varepsilon_k$.

The specific heat reads as

$$C_V = \frac{T}{N} \left(\frac{\partial S}{\partial T} \right)_{V\mu} = \frac{\beta^2}{N} \sum_k \omega_k^2 n_k (1 \pm n_k) , \quad (165)$$

where the derivative

$$\frac{\partial S}{\partial T} = \beta^3 \sum_k \omega_k^2 n_k (1 \pm n_k)$$

is used. For a stable system, it should be: $0 \leq C_V < \infty$.

For the isothermal compressibility, using the derivative

$$\frac{\partial N}{\partial \mu} = \beta \sum_k n_k (1 \pm n_k) ,$$

we have

$$\kappa_T = \frac{1}{\rho N} \left(\frac{\partial N}{\partial \mu} \right)_{TV} = \frac{\text{var}(\hat{N})}{\rho N T} , \quad (166)$$

with the number-operator dispersion

$$\text{var}(\hat{N}) = \sum_k n_k (1 \pm n_k) . \quad (167)$$

The system is stable provided that $0 \leq \kappa_T < \infty$, which implies that $\text{var}(\hat{N}) \sim N$.

Similar expressions arise for interacting systems treated in a mean-field approximation, with the difference that ω_k becomes an effective spectrum defined by the corresponding self-consistency conditions.

5.3 Correlation functions

In the case of the quadratic grand Hamiltonian (160), it is straightforward to calculate various correlation functions. Let b_i be either a_{k_i} or $a_{k_i}^\dagger$. And let us consider the correlation functions of the type $\langle b_1 b_2 \dots b_{2n} \rangle$, containing even numbers of the field operators. Such correlators can be expressed through the combinations of the lower-order correlators, as we shall illustrate for the fourth-order correlation function.

Wick theorem. The correlation function $\langle b_1 b_2 b_3 b_4 \rangle$, where the averaging is defined using the equilibrium statistical operator $\hat{\rho} = e^{-\beta H}/Z$, with Hamiltonian (160), reads as

$$\langle b_1 b_2 b_3 b_4 \rangle = \langle b_1 b_2 \rangle \langle b_3 b_4 \rangle \pm \langle b_1 b_3 \rangle \langle b_2 b_4 \rangle + \langle b_1 b_4 \rangle \langle b_2 b_3 \rangle . \quad (168)$$

Proof. With Hamiltonian (160), we have

$$a_k H = (H + \omega_k) a_k , \quad a_k H^n = (H + \omega_k)^n a_k ,$$

from where

$$a_k e^{-\beta H} = e^{-\beta H} a_k e^{-\beta \omega_k} , \quad a_k^\dagger e^{-\beta H} = e^{-\beta H} a_k^\dagger e^{\beta \omega_k} .$$

Therefore

$$(1 \pm n_k) a_k \hat{\rho} = n_k \hat{\rho} a_k , \quad n_k a_k^\dagger \hat{\rho} = \hat{\rho} a_k^\dagger (1 \pm n_k) .$$

Using this, for any operator \hat{A} , we get

$$(1 \pm n_k) \langle \hat{A} a_k \rangle = n_k \langle a_k \hat{A} \rangle , \quad n_k \langle \hat{A} a_k^\dagger \rangle = (1 \pm n_k) \langle a_k^\dagger \hat{A} \rangle ,$$

$$\langle a_k \hat{A} \rangle = (1 \pm n_k) \langle [a_k, \hat{A}]_\mp \rangle , \quad \langle a_k^\dagger \hat{A} \rangle = \mp n_k \langle [a_k^\dagger, \hat{A}]_\mp \rangle .$$

Recall the commutation relations

$$[a_k, a_p]_\mp = 0 , \quad [a_k^\dagger, a_p^\dagger]_\mp = 0 ,$$

$$[a_k, a_p^\dagger]_\mp = \delta_{kp} , \quad [a_k^\dagger, a_p]_\mp = \mp \delta_{kp}$$

and the properties

$$\langle a_k^\dagger a_p \rangle = \delta_{kp} n_k , \quad \langle a_k a_p^\dagger \rangle = \delta_{kp} (1 \pm n_k) .$$

Excluding here δ_{kp} , we find

$$(1 \pm n_k)[a_k, b_i]_\mp = \langle a_k b_i \rangle , \quad \mp n_k [a_k^\dagger, b_i]_\mp = \mp \langle a_k^\dagger b_i \rangle .$$

Using the equality

$$[A, BCD]_\mp = [A, B]_\mp CD \pm B[A, C]_\mp D + BC[A, D]_\mp ,$$

valid for any operators A, B, C, D , we come to the expression

$$\langle a_k b_2 b_3 b_4 \rangle = \langle a_k b_2 \rangle \langle b_3 b_4 \rangle \pm \langle a_k b_3 \rangle \langle b_2 b_4 \rangle + \langle a_k b_4 \rangle \langle b_2 b_3 \rangle ,$$

and to the same expression for a_k^\dagger . This proves the theorem. \square

Because of relation (156), the equivalent representation is valid for the field operators $\psi(x)$. Let ψ_i be either $\psi(x_i)$ or $\psi^\dagger(x_i)$. Then

$$\langle \psi_1 \psi_2 \psi_3 \psi_4 \rangle = \langle \psi_1 \psi_2 \rangle \langle \psi_3 \psi_4 \rangle \pm \langle \psi_1 \psi_3 \rangle \langle \psi_2 \psi_4 \rangle + \langle \psi_1 \psi_4 \rangle \langle \psi_2 \psi_3 \rangle .$$

Generally, any correlator $\langle b_1 b_2 \dots b_{2n} \rangle$, calculated with the quadratic Hamiltonian (160), results in the combination

$$\begin{aligned} \langle b_1 b_2 b_3 b_4 \dots b_{2n} \rangle &= \langle b_1 b_2 \rangle \langle b_3 b_4 \rangle \dots \langle b_{2n-1} b_{2n} \rangle \pm \\ &\pm \langle b_1 b_3 \rangle \langle b_2 b_4 \rangle \dots \langle b_{2n-1} b_{2n} \rangle + \langle b_1 b_4 \rangle \langle b_2 b_3 \rangle \dots \langle b_{2n-1} b_{2n} \rangle + \dots . \end{aligned}$$

In the same way, the correlator $\langle \psi_1 \psi_2 \dots \psi_{2n} \rangle$ can be represented as the combination of the products of all possible pair *contractions* $\langle \psi_i \psi_j \rangle$.

However, one has to be cautious when the quadratic Hamiltonian is not exact, but is a result of an approximation. Then calculating the correlators of the orders higher than second can lead to unreasonable conclusions. This is the simple consequence of the approximation theory prescribing that, if the main Hamiltonian is of second order, then only the correlators of second order are reliable. In brief, accepting an approximation of some order, one should not go outside of the region of applicability of that approximation, but must restrain calculations to the same order.

5.4 Canonical transformations

The general form of a quadratic Hamiltonian is

$$H = \int \left[\psi^\dagger(x) A(x, x') \psi(x') + \frac{1}{2} B(x, x') \psi^\dagger(x) \psi^\dagger(x') + \frac{1}{2} B^*(x, x') \psi(x) \psi(x') \right] dx dx' . \quad (169)$$

Such a quadratic form can be diagonalized by means of canonical transformations [30].

The Hamiltonian can be rewritten in the quantum-number representation, for which one expands the field operators over an orthonormal complete basis, labeled by quantum multi-indices k ,

$$\psi(x) = \sum_k a_k \varphi_k(x) , \quad a_k = \int \varphi_k^*(x) \psi(x) dx .$$

Introducing the notation

$$A(x, x') \equiv \sum_{kp} A_{kp} \varphi_k(x) \varphi_p^*(x') , \quad B(x, x') \equiv \sum_{kp} B_{kp} \varphi_k(x) \varphi_p(x') ,$$

one gets the Hamiltonian

$$H = \sum_{kp} \left(A_{kp} a_k^\dagger a_p + \frac{1}{2} B_{kp} a_k^\dagger a_p^\dagger + \frac{1}{2} B_{kp}^* a_p a_k \right) . \quad (170)$$

In general, the Hamiltonian can contain nonoperator terms, which are not included here.

The Hamiltonian is self-adjoint, $H^\dagger = H$, which requires that $A_{kp}^* = A_{pk}$. Also, it should be that $B_{pk} = \pm B_{kp}$, since

$$\sum_{kp} B_{kp} a_k^\dagger a_p^\dagger = \pm \sum_{kp} B_{kp} a_p^\dagger a_k^\dagger = \pm \sum_{kp} B_{pk} a_k^\dagger a_p^\dagger .$$

The general form of the *Bogolubov canonical transformations* is

$$a_k = \sum_p (u_{kp} b_p + v_{kp}^* b_p^\dagger) , \quad a_k^\dagger = \sum_p (u_{kp}^* b_p^\dagger + v_{kp} b_p) . \quad (171)$$

It is termed canonical because it does not change the particle statistics, so that the commutation relations for the operators a_k are the same as these for the operators b_k :

$$[a_k, a_p]_\mp = 0 , \quad [a_k, a_p^\dagger]_\mp = \delta_{kp} ,$$

$$[b_k, b_p]_\mp = 0 , \quad [b_k, b_p^\dagger]_\mp = \delta_{kp} .$$

This imposes the restrictions

$$\sum_q (u_{kq}^* v_{pq} \mp v_{kq} u_{pq}^*) = 0 , \quad \sum_q (u_{kq}^* u_{pq} \mp v_{kq} v_{pq}^*) = \delta_{kp} .$$

One requires that the Hamiltonian would acquire the diagonal form called the *Bogolubov Hamiltonian*

$$H = \sum_k \omega_k b_k^\dagger b_k + E_0 . \quad (172)$$

The diagonalization condition gives

$$E_0 = E_0^* = - \sum_{kp} \omega_k |v_{pk}|^2 ,$$

while the spectrum $\omega_k = \omega_k^*$ and the expansion functions are defined by the *Bogolubov equations*

$$\sum_q [(A_{kq} - \delta_{kq}\omega_p)u_{qp} + B_{kq}v_{qp}] = 0, \quad \sum_q [(A_{kq}^* + \delta_{kq}\omega_p)v_{qp} + B_{kq}^*u_{qp}] = 0. \quad (173)$$

The operators a_k describe particles, while the operators b_k correspond to quasiparticles, that is, to excitations. The distribution of particles is

$$n_k \equiv \langle a_k^\dagger a_k \rangle = \sum_p [(|u_{kp}|^2 \pm |v_{kp}|^2) \pi_k + |v_{kp}|^2],$$

while the distribution of excitations is

$$\pi_k \equiv \langle b_k^\dagger b_k \rangle = \frac{1}{e^{\beta\omega_k} \mp 1}.$$

The particle distribution is summed to the total number of particles $N = \sum_k n_k$. But the sum over the excitation distribution is not fixed, being dependent on thermodynamic variables.

Because of the diagonal form of the Bogolubov Hamiltonian (172), one has

$$\langle b_k \rangle = 0, \quad \langle b_k b_p \rangle = 0, \quad \langle b_k^\dagger b_p \rangle = \delta_{kp} \pi_k, \quad \langle b_k b_p^\dagger \rangle = \delta_{kp} (1 \pm \pi_k).$$

As follows from transformation (171), $\langle a_k^\dagger \rangle = 0$, that is, there is no generation of single particles. However, there can exist the anomalous average

$$\langle a_k a_p \rangle = \sum_q [(v_{kq}^* u_{pq} \pm u_{kq} v_{pq}^*) \pi_q + u_{kq} v_{pq}^*].$$

The nonzero $\langle a_k a_p \rangle \neq 0$ implies the breaking of gauge symmetry for pairs, which can be created or annihilated together. Respectively, the anomalous average

$$\langle \psi(x) \psi(x') \rangle = \sum_{kp} \langle a_k a_p \rangle \varphi_k(x) \varphi_p(x')$$

is also nonzero.

5.5 Linear terms

It may happen that the Hamiltonian contains the terms linear in the field operators, as in the expression

$$H = \sum_{kp} \left(A_{kp} a_k^\dagger a_p + \frac{1}{2} B_{kp} a_k^\dagger a_p^\dagger + \frac{1}{2} B_{kp}^* a_p a_k \right) + \sum_k \left(C_k a_k^\dagger + C_k^* a_k \right). \quad (174)$$

This Hamiltonian is not gauge invariant, because of which $\langle a_k \rangle$ is, generally, nonzero, meaning that these particles can be created and annihilated. Since the particle spin is conserved, such a situation can occur only for particles of spin 0, that is, for Bose particles. Therefore, the operators a_k here satisfy the Bose commutation relations.

To reduce Hamiltonian (174) to a diagonal form, one, first, diagonalizes the quadratic part of the Hamiltonian as earlier in the previous section, which yields

$$H = \sum_k \left(\omega_k b_k^\dagger b_k + D_k b_k^\dagger + D_k^* b_k \right) + E_0 ,$$

with

$$D_k \equiv \sum_p \left(C_p u_{pk}^* + C_p^* v_{pk} \right) .$$

Then one invokes another canonical transformation

$$\hat{b}_k = b_k + \frac{D_k}{\omega_k} , \quad (175)$$

where the operators \hat{b}_k satisfy the same Bose commutation relations as the operators a_k ,

$$[\hat{b}_k, \hat{b}_p] = 0 , \quad [\hat{b}_k, \hat{b}_p^\dagger] = \delta_{kp} .$$

As a result, we come to the diagonal form

$$H = \sum_k \omega_k \hat{b}_k^\dagger \hat{b}_k + \tilde{E}_0 , \quad (176)$$

in which

$$\tilde{E}_0 = E_0 - \sum_k \frac{|D_k|^2}{\omega_k} .$$

The diagonal form of Hamiltonian (176) tells us that

$$\langle \hat{b}_k \rangle = 0 , \quad \langle \hat{b}_k \hat{b}_p \rangle = 0 ,$$

and the distribution of excitations is given by the equation

$$\langle \hat{b}_k^\dagger \hat{b}_p \rangle = \frac{\delta_{kp}}{e^{\beta \omega_k} \mp 1} .$$

However the quasiparticles, characterized by the operators b_k , are not conserved and can be created or destructed as single objects as well as pairs:

$$\langle b_k \rangle = - \frac{D_k}{\omega_k} , \quad \langle b_k b_p \rangle = \frac{D_k D_p}{\omega_k \omega_p} , \quad \langle b_k^\dagger b_k \rangle = \langle \hat{b}_k^\dagger \hat{b}_k \rangle + \frac{|D_k|^2}{\omega_k^2} .$$

The same concerns the original particles described by the operators a_k , which can be created or annihilated by pairs as well as by single particles. For example,

$$\langle a_k \rangle = - \sum_p \left(\frac{D_p}{\omega_p} u_{kp} + \frac{D_p^*}{\omega_p} v_{kp}^* \right) .$$

This shows that, if the Hamiltonian contains the terms linear in the field operators, as in Eq. (174), hence the gauge symmetry is broken, then $\langle a_k \rangle$ is not zero.

5.6 Hartree approximation

The Hamiltonian of interacting particles is not quadratic and, generally, is given by the expression

$$H = \int \psi^\dagger(x) \left[-\frac{\nabla^2}{2m} + U(x) - \mu \right] \psi(x) dx + \frac{1}{2} \int \psi^\dagger(x) \psi^\dagger(x') \Phi(x, x') \psi(x') \psi(x) dx dx' . \quad (177)$$

To reduce it to a quadratic form, one needs to resort to some approximations. The simplest such an approximation is the *Hartree approximation* that, for averages, corresponds to the decoupling

$$\langle \psi^\dagger(x_1) \psi^\dagger(x_2) \psi(x_3) \psi(x_4) \rangle = \langle \psi^\dagger(x_1) \psi(x_4) \rangle \langle \psi^\dagger(x_2) \psi(x_3) \rangle . \quad (178)$$

In terms of the field operators, this is equivalent to the equality

$$\begin{aligned} \psi^\dagger(x_1) \psi^\dagger(x_2) \psi(x_3) \psi(x_4) &= \psi^\dagger(x_1) \psi(x_4) \langle \psi^\dagger(x_2) \psi(x_3) \rangle + \\ &+ \langle \psi^\dagger(x_1) \psi(x_4) \rangle \psi^\dagger(x_2) \psi(x_3) - \langle \psi^\dagger(x_1) \psi(x_4) \rangle \langle \psi^\dagger(x_2) \psi(x_3) \rangle . \end{aligned} \quad (179)$$

Introducing the *Hartree potential*

$$U_H(x) \equiv \int \Phi(x, x') \rho(x') dx'$$

and the average density

$$\rho(x) = \langle \hat{n}(x) \rangle , \quad \hat{n}(x) \equiv \psi^\dagger(x) \psi(x)$$

makes it possible to reduce the initial Hamiltonian (177) to the Hartree approximation

$$H = \int \psi^\dagger(x) \left[-\frac{\nabla^2}{2m} + U(x) + U_H(x) - \mu \right] \psi(x) dx + E_0 , \quad (180)$$

where

$$E_0 = -\frac{1}{2} \int \Phi(x, x') \rho(x) \rho(x') dx dx' .$$

If we choose the orthonormal basis $\{\varphi_k(x)\}$ as the set of solutions to the equation

$$\left[-\frac{\nabla^2}{2m} + U(x) + U_H(x) \right] \varphi_k(x) = \varepsilon_k \varphi_k(x) ,$$

then, expanding the field operators $\psi(x)$ over this basis, we come to the Hamiltonian

$$H = \sum_k \omega_k a_k^\dagger a_k + E_0 \quad (\omega_k = \varepsilon_k - \mu) . \quad (181)$$

For the average density, we have

$$\rho(x) = \sum_k n_k |\varphi_k(x)|^2 , \quad n_k = \langle a_k^\dagger a_k \rangle = (e^{\beta \omega_k} \mp 1)^{-1} .$$

As an example, let us consider a uniform system, when there is no external potential, $U(\mathbf{r}) = 0$, and let $x = \mathbf{r}$. Also, we take into account that the interaction potential usually depends on the difference of the spatial variables, $\Phi(\mathbf{r}, \mathbf{r}') = \Phi(\mathbf{r} - \mathbf{r}')$. Then the expansion basis is composed of the plane waves $\varphi_k(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}/\sqrt{V}$ and the average density is

$$\rho(\mathbf{r}) = \frac{1}{V} \sum_k n_k = \rho ,$$

which is a constant defining the chemical potential $\mu = \mu(T, \rho)$. Consequently,

$$U_H(\mathbf{r}) = \rho\Phi_0 , \quad \Phi_0 \equiv \int \Phi(\mathbf{r}) d\mathbf{r} , \quad \varepsilon_k = \frac{k^2}{2m} + \rho\Phi_0 .$$

The internal energy of the system is

$$E = \langle H \rangle + \mu N = E_0 + \sum_k n_k \varepsilon_k , \quad E_0 = -\frac{1}{2} N \rho \Phi_0 ,$$

with E_0 being the ground-state energy.

As is seen, for a uniform system, the Hartree approximation is applicable only if the interaction potential is integrable:

$$\left| \int \Phi(\mathbf{r}) d\mathbf{r} \right| < \infty .$$

Another example is a periodic system with a crystalline lattice defined by the crystalline vectors $\{\mathbf{a}_i\}$, with $i = 1, 2, \dots, N$ enumerating the lattice sites. Then the expansion basis can be formed by well localized functions $\varphi_i(\mathbf{r}) = \varphi(\mathbf{r} - \mathbf{a}_i)$, called localized orbitals. Good localization means that the *overlap integral*

$$(\varphi_i, \varphi_j) \equiv \int \varphi^*(\mathbf{r} - \mathbf{a}_i) \varphi(\mathbf{r} - \mathbf{a}_j) d\mathbf{r}$$

is small. By the Cauchy-Schwarz inequality,

$$|(\varphi_i, \varphi_j)|^2 \leq |(\varphi_i, \varphi_i)(\varphi_j, \varphi_j)| .$$

Since the orbitals are normalized, $(\varphi_i, \varphi_i) = 1$, one has $|(\varphi_i, \varphi_j)| \leq 1$. And the orbitals are well localized, when $|(\varphi_i, \varphi_j)| \ll 1$ ($i \neq j$). If the orbitals are so well localized that they do not intersect in space, then the requirement for the integrability of the interaction potential is not compulsory.

5.7 Hartree-Fock approximation

The Hartree approximation can work well when the particle exchange is negligible, for instance, as in the case of well localized orbitals. When the particle exchange is important, one has to resort to the Hartree-Fock approximation. This approximation implies for the averages the decoupling

$$\langle \psi^\dagger(x_1) \psi^\dagger(x_2) \psi(x_3) \psi(x_4) \rangle = \langle \psi^\dagger(x_1) \psi(x_4) \rangle \langle \psi^\dagger(x_2) \psi(x_3) \rangle \pm .$$

$$\pm < \psi^\dagger(x_1)\psi(x_3) > < \psi^\dagger(x_2)\psi(x_4) > . \quad (182)$$

And in terms of the operators, this is equivalent to the construction

$$\begin{aligned} \psi^\dagger(x_1)\psi^\dagger(x_2)\psi(x_3)\psi(x_4) &= \psi^\dagger(x_1)\psi(x_4) < \psi^\dagger(x_2)\psi(x_3) > + \\ &+ < \psi^\dagger(x_1)\psi(x_4) > \psi^\dagger(x_2)\psi(x_3) - < \psi^\dagger(x_1)\psi(x_4) > < \psi^\dagger(x_2)\psi(x_3) > \pm \\ &\pm \psi^\dagger(x_1)\psi(x_3) < \psi^\dagger(x_2)\psi(x_4) > \pm < \psi^\dagger(x_1)\psi(x_3) > \psi^\dagger(x_2)\psi(x_4) \mp \\ &\mp < \psi^\dagger(x_1)\psi(x_3) > < \psi^\dagger(x_2)\psi(x_4) > . \end{aligned} \quad (183)$$

Let us define the Hartree potential $U_H(x)$, as in the previous section. And let us introduce the *Hartree-Fock operator potential*

$$\hat{U}_{HF}(x) \equiv U_H(x) + \hat{U}_{ex}(x) ,$$

where $\hat{U}_{ex}(x)$ is the *exchange operator potential* defined by the action

$$\hat{U}_{ex}(x)\psi(x) \equiv \pm \int \Phi(x, x')\rho_1(x, x')\psi(x') dx' .$$

Then the Hartree-Fock approximation for Hamiltonian (177) leads to

$$H = \int \psi^\dagger(x) \left[-\frac{\nabla^2}{2m} + U(x) + \hat{U}_{HF}(x) - \mu \right] \psi(x) dx + E_0^{HF} , \quad (184)$$

where

$$E_0^{HF} = -\frac{1}{2} \int \Phi(x, x') [\rho(x)\rho(x') \pm |\rho_1(x, x')|^2] dx dx'$$

and the first-order density matrix is

$$\rho_1(x, x') = < \psi^\dagger(x')\psi(x) > = \sum_k n_k \varphi_k(x) \varphi_k^*(x') .$$

One can choose the expansion basis made of the functions that are the solutions to the equation

$$\left[-\frac{\nabla^2}{2m} + U(x) + \hat{U}_{HF}(x) \right] \varphi_k(x) = \varepsilon_k \varphi_k(x) .$$

The difference with the Hartree approximation is that now $\hat{U}_{HF}(x)$ is a nonlocal potential.

For a uniform system, when $x \rightarrow \mathbf{r}$ and $U(\mathbf{r}) = 0$, the expansion basis is composed of the plane waves. The exchange operator potential acts on the latter according to the rule

$$\hat{U}_{ex}(\mathbf{r})\varphi_k(\mathbf{r}) = \pm \int \Phi(\mathbf{r} - \mathbf{r}')\rho_1(\mathbf{r}, \mathbf{r}')\varphi_k(\mathbf{r}') d\mathbf{r}' .$$

Assuming that the interaction potential is integrable, one invokes the Fourier transform

$$\Phi(\mathbf{r}) = \frac{1}{V} \sum_k \Phi_k e^{i\mathbf{k}\cdot\mathbf{r}} , \quad \Phi_k = \int \Phi(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{r}} d\mathbf{r} .$$

Hamiltonian (184) becomes

$$H = \sum_k (\varepsilon_k - \mu) a_k^\dagger a_k + E_0^{HF} , \quad (185)$$

with the spectrum

$$\varepsilon_k = \frac{k^2}{2m} + \rho\Phi_0 \pm \frac{1}{V} \sum_p n_p \Phi_{k-p}$$

and the nonoperator term

$$E_0^{HF} = -\frac{1}{2} \left(N\rho\Phi_0 \pm \frac{1}{V} \sum_{kp} \Phi_{k+p} n_k n_p \right) .$$

In the case of an isotropic system, one has $\Phi_{-k} = \Phi_k$ and $n_{-k} = n_k$.

5.8 Hartree-Fock-Bogolubov approximation

The Hartree-Fock approximation preserves the gauge symmetry. Therefore it cannot be applied for systems where gauge symmetry becomes broken. That is, the Hartree-Fock approximation is not applicable for systems with Bose-Einstein condensate as well as for superconducting systems.

The approximation that takes into account possible gauge symmetry breaking is the Hartree-Fock-Bogolubov approximation that for the averages reads as the decoupling

$$\begin{aligned} < \psi^\dagger(x_1) \psi^\dagger(x_2) \psi(x_3) \psi(x_4) > = < \psi^\dagger(x_1) \psi(x_4) > < \psi^\dagger(x_2) \psi(x_3) > \pm \\ & \pm < \psi^\dagger(x_1) \psi(x_3) > < \psi^\dagger(x_2) \psi(x_4) > + < \psi^\dagger(x_1) \psi^\dagger(x_2) > < \psi(x_3) \psi(x_4) > . \end{aligned} \quad (186)$$

This differs from the Hartree-Fock approximation by the term containing the anomalous averages $< \psi(x) \psi(x') >$ and $< \psi^\dagger(x) \psi^\dagger(x') >$. To express the Hartree-Fock-Bogolubov approximation in the operator terms, we shall use the short-hand notation for the field operators $\psi_i \equiv \psi(x_i)$ and $\psi_i^\dagger \equiv \psi^\dagger(x_i)$. Then this approximation corresponds to the operator expression

$$\begin{aligned} \psi_1^\dagger \psi_2^\dagger \psi_3 \psi_4 &= \psi_1^\dagger \psi_4 < \psi_2^\dagger \psi_3 > + < \psi_1^\dagger \psi_4 > \psi_2^\dagger \psi_3 - \\ &- < \psi_1^\dagger \psi_4 > < \psi_2^\dagger \psi_3 > \pm \psi_1^\dagger \psi_3 < \psi_2^\dagger \psi_4 > \pm < \psi_1^\dagger \psi_3 > \psi_2^\dagger \psi_4 \mp \\ &\mp < \psi_1^\dagger \psi_3 > < \psi_2^\dagger \psi_4 > + \psi_1^\dagger \psi_2^\dagger < \psi_3 \psi_4 > + < \psi_1^\dagger \psi_2^\dagger > \psi_3 \psi_4 - < \psi_1^\dagger \psi_2^\dagger > < \psi_3 \psi_4 > . \end{aligned} \quad (187)$$

Hamiltonian (177) reduces to the form

$$\begin{aligned} H &= \int \psi^\dagger(x) \left[-\frac{\nabla^2}{2m} + U(x) + \hat{U}_{HF}(x) - \mu \right] \psi(x) dx + \\ &+ \frac{1}{2} \int \Phi(x, x') \left[\psi^\dagger(x) \psi^\dagger(x') < \psi(x') \psi(x) > + \psi(x') \psi(x) < \psi^\dagger(x) \psi^\dagger(x') > \right] dx dx' + E_0^{HFB} , \end{aligned} \quad (188)$$

in which

$$E_0^{HFB} = E_0^{HF} - \frac{1}{2} \int \Phi(x, x') | < \psi(x') \psi(x) > |^2 dx dx' .$$

Expanding the field operators over a basis and introducing the notation

$$A_{kp} \equiv \int \varphi_k^*(x) \left[-\frac{\nabla^2}{2m} + U(x) + \hat{U}_{HF}(x) - \mu \right] \varphi_p(x) dx ,$$

$$B_{kp} \equiv \int \varphi_k^*(x) \Phi(x, x') < \psi(x') \psi(x) > \varphi_p^*(x') dx dx' ,$$

we can transform Hamiltonian (188) to the expression

$$H = \sum_{kp} \left(A_{kp} a_k^\dagger a_p + \frac{1}{2} B_{kp} a_k^\dagger a_p^\dagger + \frac{1}{2} B_{kp}^* a_p a_k \right) + E_0^{HFB} . \quad (189)$$

The operator part here is equivalent to that in Eq. (170), hence can be diagonalized in the same way as in Sec. 5.4.

As in the previous cases, the interaction potential, generally, has to be integrable for the applicability of this approximation.

5.9 Kirkwood approximation

Quite often, the interaction potentials are not integrable, because of which the direct application of the mean-field approximations of the previous sections can be inadmissible.

An example of a nonintegrable potential is the singular power-law potential $\Phi(\mathbf{r}) \propto 1/r^n$, where $r \equiv |\mathbf{r}|$. In the three-dimensional space, the integral

$$\int \Phi(\mathbf{r}) d\mathbf{r} \propto \int_0^R \frac{dr}{r^{n-2}} ,$$

where R is the system radius, diverges at zero, if $n \geq 3$. This is classified as ultraviolet divergence.

A rather popular interaction potential is the Lennard-Jones potential

$$\Phi(\mathbf{r}) = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right] ,$$

describing the interactions between many neutral particles, such as atoms and molecules. The *nonintegrable* potentials are often termed the *hard-core* potentials. In the case of the Lennard-Jones potential, σ is the hard-core radius.

Kirkwood [47] suggested a method of dealing with hard-core interaction potentials by employing the modified Hartree decoupling

$$< \psi^\dagger(\mathbf{r}) \psi^\dagger(\mathbf{r}') \psi(\mathbf{r}') \psi(\mathbf{r}) > = g(\mathbf{r}, \mathbf{r}') < \psi^\dagger(\mathbf{r}) \psi(\mathbf{r}) > < \psi^\dagger(\mathbf{r}') \psi(\mathbf{r}') > , \quad (190)$$

which differs from the Hartree approximation by the factor $g(\mathbf{r}, \mathbf{r}')$ that takes into account short-range particle correlations, tending to zero, where the interaction potential tends to infinity. The actual form of $g(\mathbf{r}, \mathbf{r}')$ should be defined from other physical assumptions. In the operator representation, the Kirkwood approximation reads as

$$\psi^\dagger(\mathbf{r}) \psi^\dagger(\mathbf{r}') \psi(\mathbf{r}') \psi(\mathbf{r}) = g(\mathbf{r}, \mathbf{r}') [\psi^\dagger(\mathbf{r}) \psi(\mathbf{r}) < \psi^\dagger(\mathbf{r}') \psi(\mathbf{r}') > +$$

$$+ < \psi^\dagger(\mathbf{r})\psi(\mathbf{r}) > \psi^\dagger(\mathbf{r}')\psi(\mathbf{r}') - < \psi^\dagger(\mathbf{r})\psi(\mathbf{r}) > < \psi^\dagger(\mathbf{r}')\psi(\mathbf{r}') >] . \quad (191)$$

Substituting form (191) into the Hamiltonian, one meets the effective *Kirkwood potential*

$$U_K(\mathbf{r}) \equiv \int \Phi(\mathbf{r}, \mathbf{r}') g(\mathbf{r}, \mathbf{r}') \rho(\mathbf{r}') d\mathbf{r}' ,$$

which is similar to the Hartree potential, but with the principal difference that now, instead of the bare interaction potential, one has the *smoothed potential*

$$\bar{\Phi}(\mathbf{r}, \mathbf{r}') \equiv g(\mathbf{r}, \mathbf{r}') \Phi(\mathbf{r}, \mathbf{r}') ,$$

The latter is integrable,

$$\left| \int \bar{\Phi}(\mathbf{r}, \mathbf{r}') d\mathbf{r}' \right| < \infty ,$$

even if the bare interaction potential is not.

Equation (190), as is evident, is nothing but the definition of the pair correlation function $g(\mathbf{r}, \mathbf{r}')$. Kirkwood suggested to extract the pair correlation function $g(\mathbf{r}, \mathbf{r}')$ from experiment. This could be done by measuring the *structure factor*

$$S(\mathbf{k}) \equiv \frac{1}{N} \int [< \hat{n}(\mathbf{r})\hat{n}(\mathbf{r}') > - \rho(\mathbf{r})\rho(\mathbf{r}')] e^{-i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')} d\mathbf{r}d\mathbf{r}' , \quad (192)$$

where $\hat{n}(\mathbf{r}) = \psi^\dagger(\mathbf{r})\psi(\mathbf{r})$ is the density operator. The density-density correlation function can be expressed through the pair correlation function:

$$< \hat{n}(\mathbf{r})\hat{n}(\mathbf{r}') > = \rho(\mathbf{r})\delta(\mathbf{r} - \mathbf{r}') + \rho(\mathbf{r})\rho(\mathbf{r}')g(\mathbf{r}, \mathbf{r}') , \quad (193)$$

which gives

$$S(\mathbf{k}) = 1 + \frac{1}{N} \int \rho(\mathbf{r})\rho(\mathbf{r}') [g(\mathbf{r}, \mathbf{r}') - 1] e^{-i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')} d\mathbf{r}d\mathbf{r}' . \quad (194)$$

In the uniform case, when $\rho(\mathbf{r}) = \rho$ and $g(\mathbf{r}, \mathbf{r}') = g(\mathbf{r} - \mathbf{r}')$, one has

$$S(\mathbf{k}) = 1 + \rho \int [g(\mathbf{r}) - 1] e^{-i\mathbf{k}\cdot\mathbf{r}} d\mathbf{r} . \quad (195)$$

Taking the inverse Fourier transform yields

$$g(\mathbf{r}) = 1 + \frac{1}{N} \sum_{\mathbf{k}} [S(\mathbf{k}) - 1] e^{i\mathbf{k}\cdot\mathbf{r}} .$$

For a large system, one passes from summation to integration, obtaining

$$g(\mathbf{r}) = 1 + \frac{1}{\rho} \int [S(\mathbf{k}) - 1] e^{i\mathbf{k}\cdot\mathbf{r}} \frac{d\mathbf{k}}{(2\pi)^3} . \quad (196)$$

Measuring in experiment the structure factor $S(\mathbf{k})$, one can calculate the pair correlation function by means of relation (196).

5.10 Correlated approximations

The Kirkwood idea can be generalized as follows [60]. For simplicity, we again use the notation of Sec. 5.8 denoting the field operators as $\psi_i \equiv \psi(\mathbf{r}_i)$. And let us employ a brief notation $g_{12} \equiv g(\mathbf{r}_1, \mathbf{r}_2)$ for the pair correlation function. The general form of a correlated approximation is

$$\psi_1^\dagger \psi_2^\dagger \psi_3 \psi_4 = \sqrt{g_{12}g_{34}} \left[\psi_1^\dagger \psi_2^\dagger \psi_3 \psi_4 \right]_{MF} , \quad (197)$$

where the term in the square brackets in the right-hand side implies some of the mean-field approximations. For instance, using for $[\dots]_{MF}$ the Hartree approximation leads to the Kirkwood approximation (190). Employing for $[\dots]_{MF}$ the Hartree-Fock approximation results in the correlated Hartree-Fock form

$$\begin{aligned} \psi^\dagger(\mathbf{r})\psi^\dagger(\mathbf{r}')\psi(\mathbf{r}')\psi(\mathbf{r}) &= g(\mathbf{r}, \mathbf{r}') \left[\psi^\dagger(\mathbf{r})\psi(\mathbf{r}) < \psi^\dagger(\mathbf{r}')\psi(\mathbf{r}') > + \right. \\ &+ < \psi^\dagger(\mathbf{r})\psi(\mathbf{r}) > \psi^\dagger(\mathbf{r}')\psi(\mathbf{r}') - < \psi^\dagger(\mathbf{r})\psi(\mathbf{r}) > < \psi^\dagger(\mathbf{r}')\psi(\mathbf{r}') > \pm \\ &\pm \psi^\dagger(\mathbf{r})\psi(\mathbf{r}') < \psi^\dagger(\mathbf{r}')\psi(\mathbf{r}) > \pm < \psi^\dagger(\mathbf{r})\psi(\mathbf{r}') > \psi^\dagger(\mathbf{r}')\psi(\mathbf{r}) \mp \\ &\left. \mp < \psi^\dagger(\mathbf{r})\psi(\mathbf{r}') > < \psi^\dagger(\mathbf{r}')\psi(\mathbf{r}) > \right] . \end{aligned} \quad (198)$$

In the same way, one can introduce the correlated Hartree-Fock-Bogolubov approximation.

One can remember that the pair interaction potentials depend on the difference between the spatial variables, so that $\Phi(\mathbf{r}, \mathbf{r}') = \Phi(\mathbf{r} - \mathbf{r}')$. Similarly, the pair correlation function also usually depends on this difference, $g(\mathbf{r}, \mathbf{r}') = g(\mathbf{r} - \mathbf{r}')$. Thus the smoothed interaction potential becomes $\bar{\Phi}(\mathbf{r}) \equiv g(\mathbf{r})\Phi(\mathbf{r})$. Since $g(\mathbf{r}) \rightarrow 0$ when $\Phi(\mathbf{r}) \rightarrow \infty$, the smoothed potential is integrable,

$$\left| \int \bar{\Phi}(\mathbf{r}) d\mathbf{r} \right| < \infty .$$

The pair correlation function is real-valued and symmetric, such that $g(-\mathbf{r}) = g(\mathbf{r})$. It can be expressed through the second-order density matrix as in Sec. 3.10. Thus, for a uniform system, it is

$$g(\mathbf{r} - \mathbf{r}') = \frac{\rho_2(\mathbf{r}, \mathbf{r}', \mathbf{r}, \mathbf{r}')}{\rho^2} . \quad (199)$$

The behavior of the pair correlation function is such that $g(\mathbf{r}) \rightarrow 1$, when $\Phi(\mathbf{r}) \rightarrow 0$. In view of the normalization

$$\int \rho_2(\mathbf{r}, \mathbf{r}', \mathbf{r}, \mathbf{r}') d\mathbf{r}d\mathbf{r}' = N(N-1) ,$$

the property

$$\frac{1}{V} \int g(\mathbf{r}) d\mathbf{r} = 1 - \frac{1}{N} \quad (200)$$

follows.

Instead of finding an approximation for the second-order density matrix, thus obtaining the pair correlation function (199), one often directly invokes some approximations for the latter. The simplest such a way is to use the *cutoff approximation*

$$g(\mathbf{r}) = \begin{cases} 0, & r \leq \sigma \\ 1, & r > \sigma . \end{cases}$$

Sometimes, one resorts to the *classical approximation*

$$g(\mathbf{r}) = \exp\{-\beta\Phi(\mathbf{r})\} ,$$

which can be used only for high temperatures.

For quantum systems, it is possible to use the *quantum approximation*, by setting $g(\mathbf{r}) = |\varphi(\mathbf{r})|^2$, with the function $\varphi(\mathbf{r})$ being defined by the Schrödinger equation for the relative motion:

$$\left[-\frac{\nabla^2}{2m_0} + \Phi(\mathbf{r}) + E \right] \varphi(\mathbf{r}) = 0 ,$$

where

$$m_0 \equiv \frac{m_1 m_2}{m_1 + m_2} , \quad m_1 = m_2 = m , \quad m_0 = \frac{m}{2} .$$

If at short distance $r \rightarrow 0$, the interaction potential behaves as

$$\Phi(\mathbf{r}) \simeq 4\varepsilon \left(\frac{\sigma}{r} \right)^{12} \quad (r \rightarrow 0) ,$$

then the solution to the above equation yields

$$\varphi(\mathbf{r}) \sim \exp \left\{ -\frac{\kappa}{2} \left(\frac{\sigma}{r} \right)^5 \right\} \quad (\mathbf{r} \rightarrow 0) ,$$

with the *correlation parameter* $\kappa \equiv 4\sqrt{m\varepsilon\sigma^2}/5$. Therefore

$$g(\mathbf{r}) \approx \exp \left\{ -\kappa \left(\frac{\sigma}{r} \right)^5 \right\} .$$

One also uses variational methods for constructing the pair correlation functions and there exist other more complicated ways of finding their approximate expressions [46]. When correlations and interactions between particles are strong, it may happen that even integrable potentials require to take account of the pair correlation function.

Different regimes of describing the system can be classified by considering *characteristic lengths*. The *interaction radius* is given by the expression

$$r_0 \equiv \sigma + \frac{\int_{r>\sigma} |\mathbf{r}| \Phi(\mathbf{r}) d\mathbf{r}}{\int_{r>\sigma} \Phi(\mathbf{r}) d\mathbf{r}} . \quad (201)$$

The *scattering length*

$$a_s \equiv \frac{m}{4\pi} \int \bar{\Phi}(\mathbf{r}) d\mathbf{r} \quad (202)$$

characterizes the intensity of interactions. These lengths should be compared with the *mean interparticle distance* a and the *thermal wavelength* λ_T ,

$$a \equiv \rho^{-1/3} , \quad \lambda_T \equiv \sqrt{\frac{2\pi}{mT}} .$$

Particle interactions are strong, when $a_s \gg a$, while interaction correlations are strong, if $r_0 \gg a$. For $\lambda_T \ll a$, one has the classical regime. And the *quantum regime* develops when $\lambda_T \sim a$, becoming essentially quantum for $\lambda_T > a$.

The approximations of the mean-field type provide sufficiently accurate description, when the particles in the system are strongly correlated, so that the influence of each of them is extended onto many other particles. In such a case, each particle can be imagined to be subject to a mean field created by other particles. Generally, there exist two kinds of correlations, interaction correlations and coherence correlations.

Interaction correlations are strong, when the interaction radius r_0 is much larger than the interatomic distance a . Then each particle experiences the action of many other particles of the system. Such interactions are termed long-range.

Coherence correlations characterize the particle correlations that develop not merely because of particle interactions but also due to particle quantum statistics. The *coherence radius* is defined as

$$r_{coh} \equiv \frac{\int r |C(\mathbf{r}, 0)| d\mathbf{r}}{\int |C(\mathbf{r}, 0)| d\mathbf{r}}, \quad (203)$$

where the correlation function

$$C(\mathbf{r}, \mathbf{r}') \equiv \frac{\rho_1(\mathbf{r}, \mathbf{r}')}{\sqrt{\rho(\mathbf{r})\rho(\mathbf{r}')}} \quad (204)$$

is expressed through the density matrix $\rho_1(\mathbf{r}, \mathbf{r}')$. The coherence correlations are strong, if the coherence radius r_{coh} is much larger than the interatomic distance a .

Summarizing, particle correlations in the system are strong, provided that at least one of the conditions holds true, either the interaction correlations or the coherence correlations are strong, so that either $r_0 \gg a$ or $r_{coh} \gg a$. That is, particle correlations are strong, when at least one of the following parameters is large:

$$\frac{r_0}{a} \gg 1, \quad \frac{r_{coh}}{a} \gg 1. \quad (205)$$

In that case, it is admissible to resort to a mean-field type approximation.

In conclusion, this Tutorial presents the basic notions of quantum statistical physics that are necessary for the correct description of quantum atomic systems. Accurately applying these notions will help to avoid mistakes that, unfortunately, are rather common in the current literature. The exposition in the Tutorial has been aimed to be, from one side, sufficiently detailed to be easily read and, from another side, quite brief, omitting lengthy discussions that could be found in the cited references. In the following parts, the material of this Tutorial will be applied for a more concrete description of cold atoms satisfying Bose-Einstein and Fermi-Dirac statistics.

Acknowledgement

The author acknowledges financial support from the Russian Foundation for Basic Research. The help from E.P. Yukalova is appreciated.

References

- [1] Lieb E H, Seiringer R, Solovej J P, Yngvason J 2005 *The Mathematics of the Bose Gas and Its Condensation* (Basel: Birkhauser)
- [2] Pitaevskii L and Stringari S 2003 *Bose-Einstein Condensation* (Oxford: Clarendon)
- [3] Letokhov V 2007 *Laser Control of Atoms and Molecules* (New York: Oxford University)
- [4] Pethick C J and Smith H 2008 *Bose-Einstein Condensation in Dilute Gases* (Cambridge: Cambridge University)
- [5] Courteille P W, Bagnato V S and Yukalov V I 2001 *Laser Phys.* **11** 659
- [6] Andersen J O 2004 *Rev. Mod. Phys.* **76** 599
- [7] Yukalov V I 2004 *Laser Phys. Lett.* **1** 435
- [8] Bongs K and Sengstock K 2004 *Rep. Prog. Phys.* **67** 907
- [9] Yukalov V I and Girardeau M D 2005 *Laser Phys. Lett.* **2** 375
- [10] Morsch O and Oberthaler M 2006 *Rev. Mod. Phys.* **78** 179
- [11] Posazhennikova A 2006 *Rev. Mod. Phys.* **78** 1111
- [12] Yukalov V I 2007 *Laser Phys. Lett.* **4** 632
- [13] Proukakis N P and Jackson B 2008 *J. Phys. B* **41** 203002
- [14] Yurovsky V A, Olshanii M and Weiss D S 2008 *Adv. At. Mol. Opt. Phys.* **55** 61
- [15] Moseley C, Fialko O and Ziegler K 2008 *Ann. Phys. (Berlin)* **17** 561
- [16] Bloch I, Dalibard J and Zwierger W 2008 *Rev. Mod. Phys.* **80** 885
- [17] Ketterle W and Zwierlein M W 2008 *Riv. Nuovo Cimento* **31** 247
- [18] Giorgini S, Pitaevskii L P and Stringari S 2008 *Rev. Mod. Phys.* **80** 1215
- [19] Yukalov V I 2009 *Laser Phys.* **19** 1
- [20] Chevy F and Mora C 2010 *Rep. Prog. Phys.* **73** 112401
- [21] Yukalov V I 2011 *Phys. Part. Nucl.* **42** 460
- [22] Yukalov V I 2012 *Laser Phys.* **22** 1145
- [23] Yukalov V I 2005 *Phys. Rev. A* **72** 033608
- [24] Kubo R 1968 *Thermodynamics* (Amsterdam: North-Holland)
- [25] Coleman A J and Yukalov V I (2000) *Reduced Density Matrices* (Berlin: Springer)

- [26] Yukalov V I (2002) *Physica A* **310** 413
- [27] Pathria R K (1972) *Statistical Mechanics* (Oxford: Pergamon)
- [28] Landau L D and Lifshitz E M 1980 *Statistical Physics* (Oxford: Pergamon)
- [29] Huang K 1987 *Statistical Mechanics* (New York: Wiley)
- [30] Bogolubov N N and Bogolubov N N Jr 1994 *Introduction to Quantum Statistical Mechanics* (Lausanne: Gordon and Breach)
- [31] Yukalov V I and Shumovsky A S 1990 *Lectures on Phase Transitions* (Singapore: World Scientific)
- [32] Uzunov D I 1993 *Introduction to Theory of Critical Phenomena* (Singapore: World Scientific)
- [33] Yukalov V I 1991 *Phys. Rep.* **208** 395
- [34] Gibbs J W 1928 *Collected Works* Vol 1 (New York: Longmans)
- [35] Gibbs J W 1931 *Collected Works* Vol 2 (New York: Longmans)
- [36] Landau L D 1965 *Collected Papers* (Oxford: Pergamon)
- [37] Ginzburg V L and Landau L D 1950 *J. Exp. Theor. Phys.* **20** 1064
- [38] Ginzburg V L and Pitaevskii L P 1958 *J. Exp. Theor. Phys.* **7** 858
- [39] Ginzburg V L and Sobyenin A A 1982 *J. Low Temp. Phys.* **49** 507
- [40] Tilley D R and Tilley J 1990 *Superfluidity and Superconductivity* (Bristol: Hilger)
- [41] Berezin F A 1966 *Method of Second Quantization* (New York: Academic)
- [42] Schweber S S 1962 *Introduction to Relativistic Quantum Field Theory* (New York: Harper and Row)
- [43] Yukalov V I 2011 *Phys. Lett. A* **375** 2797
- [44] Bogolubov N N 1962 *Problems of Dynamical Theory in Statistical Physics* (Amsterdam: North-Holland)
- [45] Klauder J R and Skagerstam B S (1985) *Coherent States* (Singapore: World Scientific)
- [46] Yukalov V I (1998) *Statistical Green's Functions* (Kingston: Queen's University)
- [47] Kirkwood J G (1965) *Quantum Statistics and Cooperative Phenomena* (New York: Gordon and Breach)
- [48] Bogolubov N N (1967) *Lectures on Quantum Statistics* Vol 1 (New York: Gordon and Breach)

- [49] Bogolubov N N (1970) *Lectures on Quantum Statistics* Vol 2 (New York: Gordon and Breach)
- [50] Gibbs J W (1902) *Elementary Principles in Statistical Mechanics* (Oxford: Oxford University)
- [51] Shannon C E and Weaver W (1949) *Mathematical Theory of Communication* (Urban: University of Illinois)
- [52] Janes E T (1957) *Phys. Rev.* **106** 620
- [53] Emch G G (1972) *Algebraic Methods in Statistical Mechanics and Field Theory* (New York: Wiley)
- [54] Ter Haar D (1977) *Lectures on Selected Topics in Statistical Mechanics* (Oxford: Pergamon)
- [55] Yukalov V I (2005) *Phys. Rev. E* **72** 066119
- [56] Kleinert H (2004) *Path Integrals* (Singapore: World Scientific)
- [57] Kubo R (1957) *J. Phys. Soc. Jap.* **12** 570
- [58] Bonch-Bruевич V L and Tyablikov S V (1962) *Green Function Method in Statistical Mechanics* (Amsterdam: North-Holland)
- [59] Yukalov V I (1991) *Int. J. Mod. Phys. B* **5** 3235
- [60] Yukalov V I (1990) *Phys. Rev. A* **42** 3324